

Next steps on global warming

The world needs a better mechanism — probably a permanent institute — for assessing the consequences of the accumulation of greenhouse gases than IPCC has proved to be.

THE Second World Climate Conference at Geneva last week (see page 189) will have disappointed those who looked to it for a beginning on the negotiated regulation of greenhouse gases, but that hope was always destined to be dashed. The conference was arranged to mull over the outcome of the World Climate Programme of the World Meteorological Organization (WMO). It is an historical accident that the first reports of the Intergovernmental Panel on Climate Change (IPCC) took precedence.

The Geneva meeting was not constituted to negotiate a treaty on greenhouse gases, but merely to collect opinions on what should be done. Negotiations proper will begin only at the meeting now arranged for Washington in February. The plan is that they should be completed by June 1992, at a meeting in Brazil. Then, with luck, there may be a signing ceremony.

The Geneva meeting has nevertheless made plain what kind of treaty people will be talking about next year. Officials of the United States and the Soviet Union insisted last week that their governments are not yet ready to accept restrictions on the emission of greenhouse gases as such. The US position is that the case for restricting carbon dioxide emissions, as advanced by IPCC, is not yet clinching and that the plan to end the production of long-lived chlorofluorohydrocarbons (CFCs) by the end of the decade will in any case take the edge off the predicted global warming. The Soviet Union is in a different case; the Soviet government has too many pressing matters on its collective mind to give global warming the attention it deserves. The best hope is that the two governments will eventually sign an undertaking to limit the emission of greenhouse gases when they are persuaded of the need to do so.

Slap in the face

Is that not a cruel slap in the face for governments, those of the European Communities (EC) in particular, which have already undertaken to reduce greenhouse emissions to present levels by the end of the decade, for example? Not necessarily. The EC declaration should be a powerful encouragement of others, but not much more. The urgent question hanging over next year's negotiations is that of whether the principal developing countries, notably Brazil, China and India, which are already substantial emitters of carbon dioxide, can be persuaded to join an international convention. But recent experi-

ence with the ozone convention shows that they will not do so unless the treaty includes an explicit undertaking that the development of industrializing nations will not be impeded by a greenhouse treaty. While other potential signatories may be willing to give such an undertaking in the abstract, it is unthinkable that its form could be negotiated in the eighteen months ahead.

It goes without saying that when the time arrives for implementing such an undertaking, it may well prove necessary that industrialized countries should agree to reduce, and not merely to stabilize, their emissions of carbon dioxide so as to leave room for growth elsewhere. Then, the EC's gesture two weeks ago may seem more like a pre-emptive claim on the natural capacity of the atmosphere to absorb carbon dioxide.

But is not the time for action now, not some undefined point in the future? That is what many will be asking. It is true, of course, that if the threat of global warming is real, the eventual difficulty and cost of containment can only be increased by delay. But equally, a greenhouse convention excluding some of the major sources of greenhouse gas emission would be a nonsense.

Global burden

Global warming, however serious a threat, is nothing if not a global threat. A treaty falling short of being comprehensive would simply be a recipe by which the signatories would shoulder the global burden. Sharing that burden equitably will be the chief purpose of the negotiations that lie ahead. From the outset, as this journal has been saying perhaps *ad nauseam*, it will take longer to get that question settled than to remove the remaining scientific uncertainties about global warming. Nobody should be surprised at that, for the task is nothing less than the creation of an equitable framework for the management of the Earth's resources and different nations' claims thereon.

That the IPCC exercise has not persuaded governments to more urgent action is also no surprise. The scientific community has emerged with credit from the excitements of the past few years, and especially from the arguments by which governments have been persuaded to action over the stratospheric ozone layer. The Vienna convention of 1986 was a considerable triumph; the two Montreal Protocols thereof, in 1988 and 1989, showed that declarations of good intentions could be followed by



national self-denying ordinances. But the ozone treaties deal with simpler problems than those occasioned by the greenhouse problem, where the economic stakes are much higher and the end-point is not a simple ban on certain manufacturing processes (in the hands of an identifiable dozen or so chemical companies), but the indefinite operation of a kind of quota system on the production of carbon dioxide. Copying out the Vienna convention and entering new numbers would never make a greenhouse treaty.

A need and a challenge

The need now, and the challenge for the scientific community, is for a more persuasive framework of argument than IPCC has provided. Even the best of the three reports first circulated in June, that concerned with the assessment of the likely degree of global warming, is deficient in several ways. No doubt it accurately represents the convergence of the predictions of the computer models constructed in the past several decades to track climatic change, but it lacks the incisive discussion of the remaining physical uncertainties that would have persuaded sceptics that their arguments had been given due consideration.

And the two other documents in circulation, dealing with the impact of climatic change and the possible remedies, are respectively inadequate and naive.

Both those who consider that global warming is a real threat (and even that it may already be apparent) and the small but considerable band of sceptics, have an interest that there should be a durable mechanism for developing a rounded approximation of the truth. Climatologists should not be offended if, in the process, it is held that climatology may be too important to be left to climatologists.

Operationally, what the world needs is not a loose committee structure like that of IPCC, but a permanent (for the time being) international institute whose members would be expert enough to make accurate and telling assessments of the burgeoning literature in this field, and to produce regular, pointed and authoritative reports. The work should be directed by a steering committee recruited both for its understanding of the political difficulties that lie ahead and for its technical expertise. If this is the role for which the International Institute of Applied Systems Analysis at Vienna has been searching since its creation in the early 1970s, that would be good fortune. Otherwise, something like it should be created anew.

The task, for the scientific community, is unusual. The threat of global warming is principally the outcome of theorizing, but the alertness engendered in the past few years should soon lead to a substantial harvest of pertinent data. The need is that this should be presented to politicians in a way that is as free as possible of noise, but which is free from simplifications that may be overturned by later knowledge. One route to such opinions is that they should be thoroughly discussed in advance by techni-

cal people from a variety of disciplines.

IPCC's failure to discuss dissenting opinions, perhaps to dismiss them, was a mistake. If, for example, one such as Dr Freeman Dyson should go about the world (as he does) saying that the disappearance of half the carbon dioxide added each year to the atmosphere casts doubt on the standard picture of global warming, he deserves an answer. (It is that the computer models start with measured concentrations of atmospheric carbon dioxide, but Dyson's argument has a bearing on the long-term course of the phenomenon.) And even if (or, perhaps, especially if) the objection is that of Mr John Sinunu, at the White House in Washington, that he will believe in global warming when the modellers can tell him why there was a Little Ice Age at the beginning of the seventeenth century, a considered reply is called for. (The modellers would rightly say that there may be an infinity of initial conditions that would generate such a cold spell, but that the data surviving from the time are inadequate to tell the difference between them.)

Weight of opinion

But why spend energy in rebutting arguments like these when the weight of opinion lies with global warming? Because, of course, there are many recent illustrations of how received opinion have proved to be mistaken. It is not, after all, so long since the received opinion was that nuclear power plants were manageably safe, and then there came the Chernobyl accident. With global warming, while the general character of the phenomenon may be well described by the simplest models, it will be surprising if some of the sceptics' arguments fail to modify the standard predictions in important ways. But can it matter if an overdrawn prediction should be a spur to action that, if not needed now, would be needed at some later stage? The flaw is that remedies for global warming touch the world's capacity to generate wealth and thus well-being, with social consequences that cannot easily be predicted. Over-cautious predictions are as dangerous as those by pollyannas.

The moral is that the mechanism of persuasion now required should be such that its self-correcting attributes are plain for all to see. And the goals should not be the persuasion of all politicians to a simple common view of what the threat of global warming entails, but to a rounded understanding of the complexities of the problem with which the world is faced, and to an understanding of how it may be ameliorated. One uncomfortable feature of the arguments of the past few years is that they have been most eagerly embraced by environmental lobbyists seeking to plead the support of the scientific community for the mistaken belief that calamity is just around the corner. The goal instead must be to learn to live with the threat of global warming without throwing away the prospect of continuing beneficent change. That calls for more robust concern for the management of the atmosphere and its minor constituents than has so far been apparent. □

Level pegging for UK research funds

- Inflation likely to erode increases
- Research councils to absorb overhead costs

London

LAST week's announcement of the 1991–92 science budget for the UK Department of Education and Science (DES) caused general disappointment in the British scientific community. Total research spending is likely to shrink in real terms, unless inflation is very rapidly brought under control. But spending on higher education in 1991–92 will be higher than expected, which may help alleviate the current university funding crisis.

The DES science budget is the largest single slice of UK government support for civil science. Early next year it will be divided up between the five research councils and the grant-in-aid of the Royal Society. The £928 million announced by the new Secretary of State for Education and Science, Kenneth Clarke (who replaced John MacGregor in the latest reshuffle of cabinet ministers), compares with £897 million being spent in 1990–91, but in the tradition of UK government statistics, a direct comparison between the two figures is difficult. The £928 million includes money received from European Communities' research programmes, and is further complicated by changes in the payment of research students' tuition fees, and a transfer of money from elsewhere in the DES budget to support academic supercomputing.

Professor Denis Noble, president of the academic pressure group Save British Science, says that however the figures are massaged, they represent "at least a freeze and possibly worse".

According to Clarke, the value of the science budget in 1991–92 will be preserved in real terms, but this assumes an inflation rate of six per cent over the coming year — an optimistic forecast, given that the current rate exceeds ten per cent.

Clarke also announced that, from 1992–93, the research councils must pay for all the costs of research they support in higher education institutions, apart from academic salaries and the cost of building maintenance, heating and lighting. At present, a large proportion of the overhead costs of research council projects in the universities are met from the grants awarded to universities by the Universities Funding Council (UFC). The change was originally planned for 1991–92 (see *Nature* 343, 199; 18 January 1990), but has been delayed by a year at the suggestion of the Advisory Board for the Research Councils, to allow a "smooth and effective transfer" of funds from the UFC to the

DES science budget. To cover research project overheads, a total of £50 million will be added to the science budget for 1992–93, and £100 million for the following year.

The Committee of Directors of Polytechnics has broadly welcomed this

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Kenneth Clarke — looking optimistic change, which should allow polytechnics to compete more effectively with the universities for research council grants.

SOUTH AFRICAN EDUCATION

Cuts cause fees to rise

Cape Town

SOUTH Africa's 21 universities are to have their state subsidies cut by 3 per cent in real terms next year, at a time when inflation is running at 15 per cent per year. Five universities (Rhodes, the Rand Afrikaans University, and the universities of South Africa, Pretoria and the Witwatersrand) have already responded by announcing fee increases of between 20 and 25 per cent, and the others are expected to follow suit shortly.

According to an agreement between the government and the Committee of University Principals, the government pays for 80 per cent of the cost of each student's place at a university. But in practice, increasing student numbers have led to this portion declining since 1986. Although next year's 3 per cent decrease in state support is in line with government cuts in other budget areas, it is the first time that universities have been hit with a decrease; this year all universities received an increase of 6% on last year's subsidies.

Nevertheless, this is the third year running that subsidies to the universities will not be calculated, as they have been in the past, according to the enrolment figures two years before the funding year (see *Nature* 340, 420; 1989). This means that subsidy entitlements have been frozen on

But Professor Terrence Burlin, rector of the Polytechnic of Central London, says he is "extremely disappointed" that the change has been delayed for a year, for no obvious reason. And the transfer of funds still leaves a massive imbalance in research support between the two sectors: polytechnics receive only about £30 million a year for research from the Polytechnics and Colleges Funding Council, compared with over £700 million research support entering the universities from the UFC.

Funding for higher education is set to increase by some ten per cent in 1991–92. This will finance about three-quarters of the planned five per cent increase in university student numbers, at current teaching costs. Although universities will still have to look for economies (or consider charging fees to students) if they are to expand as planned, the UFC's budget for 1991–92 is greater than many vice-chancellors had expected. But as with the DES science budget, the increase in real terms will depend on the inflation rate over the coming year. Academic pay rises may also erode the increased budget. Clarke told the House of Commons last week that a proportion of the universities' grants may be held back in the absence of a "satisfactory" pay settlement.

Peter Aldhous

the basis of 1986 enrolment figures, which has had serious consequences for universities that have grown most since then — the universities of Zululand, the Western Cape, Fort Hare and South Africa are currently operating on subsidies which constitute less than 70 per cent of their entitlement, and will be in an even worse position next year.

Higher fees make university education less accessible to less affluent sectors of the population, and will do nothing to narrow the difference in participation rates between white and black South Africans, which now stand at 30 and 3 per thousand of population respectively. Students receive no maintenance grants from the South African state, so these costs, as well as fees, have to be met by themselves or their parents. Some receive bursaries (with service contracts) from the government or the private sector, and the balance made up by loans from banks, which require either parents or the university to stand surety for them. University of Cape Town vice-chancellor Stuart Saunders responded to the cuts by calling for the establishment of a national state bursary fund to help university students, especially from underprivileged communities.

Michael Cherry

Reforms growing on the farm

Washington

US agricultural research, long criticized for its lack of peer review and reliance on political handouts, will get an overhaul and expansion following Congressional approval of a five-year, \$1,000 million research programme. The US Department of Agriculture (USDA) intends to increase the fraction of competitive, peer-reviewed grants, similar to those awarded by the National Institutes of Health (NIH) and the National Science Foundation, from less than 5 per cent of the agency's total research, to over 30 per cent.

Known as the National Research Initiative, the programme also marks the beginning of a ten-year, \$500 million project to map the genomes of key food plants, an effort that will run in parallel with NIH's Human Genome Project.

Although the programme got strong verbal support from Congress, it emerged from last month's budget negotiations with only three-quarters — \$73 million — of its expected first-year funding, and a surprisingly restrictive 14 per cent cap on "indirect costs" at the universities that receive grants was imposed.

Of the two, the cap on overheads may be the more serious problem. Inserted into the 1991 appropriations bill by congressional defenders of the traditional agricultural schools, the provision says that universities may only take 14 per cent off the top of a USDA grant to pay for overhead costs such as heating, building upkeep, and support staff. That figure is in line with the lowest figures from agricultural schools, which tend to have lower costs, but is far below the 60–80 per cent figure common to many large research universities and institutions.

Rosemary Grady, programme director for the initiative, notes that last year congress mandated a 25 per cent cap on USDA research grants and "that didn't seem to hurt the quality or the quantity of our applications." Although many universities had to take a loss on portions of a USDA grant, only one actually turned an award down last year, she says.

But a 14 per cent cap may be another matter entirely. Grady says "there is no way to predict what the impact of that will be", given that universities have already shown that they are willing to accept a low indirect-cost ceiling in order to get some USDA grants. But Jane Corlette, director of government relations at Harvard University, says "there's obviously a limit at which that become impossible". Harvard's normal indirect cost rate is about 60 per cent. "In specific cases, if the dean is willing to take the loss, we may accept the [14 per cent] rate, but not happily. It's clearly not fair. USDA is not paying what the research costs." As a

result, she says, "we can't afford to take a lot of USDA money." Despite such misgivings, the programme represents a \$30 million increase in USDA's competitive research programme.

Funding, according to the USDA programme plan, is to rise by \$50 million a year to \$300 million in 1995.

The initiative, according to John Jordan, administrator of USDA's cooperative state research service, will be divided into four general areas of priority: 40 per cent of funding will go to individual investigator-initiated grants, modelled loosely after NIH awards; 30 per cent will go to multi-disciplinary team grants, such as those between economists and agronomists; 20 per cent will go to "mission-linked" grants, such as the plant genome programme; and 10 per cent will go to strengthening institutions and individual laboratories that have not traditionally been strong agricultural research centres.

Although the plant genome effort is scheduled to run at about \$50 million per year, Congress appropriated only about \$22 million for genome studies in 1991. With internal funding shifts, "and a good tail wind", the programme may finish the year at about \$25 million, a level that will allow the programme to start and ramp up to full funding next year, Jordan says.

Christopher Anderson

UK RESEARCH COUNCILS

Belt-tightening at MRC

London

ESTIMATED overspending by £3.5 million during the current year has forced the Medical Research Council (MRC) to freeze new staff appointments and to postpone the purchase of expensive items of equipment until the next financial year begins in April 1991. Eighty-three researchers who won grants in July have also been asked to delay the start of their projects.

The new measures are the latest in a series of cutbacks forced upon the MRC by financial pressures, and follow decisions to close several MRC research groups in order to release funds for new research in the coming years (see *Nature*, **346**, 685; 23 August 1990 & **348**, 6; 1 November 1990). Nick Winterton, head of the MRC secretariat, says the measures are necessary because both inflation and pay increases have exceeded the MRC's forecasts.

Although the research councils may carry over a small surplus or deficit from year to year, Winterton says it would be "bad financial management" not to balance the books this year, given the

Misread all about it . . .

Melbourne

MOST scientists, according to a survey conducted by the Royal Melbourne Institute of Technology (RMIT), are unhappy with the way daily newspapers cover science stories — but it reports, they feel happier when they are in the story. Of 85 scientists questioned by the RMIT, three-quarters criticized news reporting for too much emphasis on "breakthroughs" and the uniqueness of results, and 70 per cent thought that information crucial to the understanding of the work was omitted.

But when asked to evaluate stories that involved them directly, these general allegations seemed no longer to apply. Only 2 per cent complained of being misquoted; 74 per cent found no inaccuracies, 88 per cent believed that no relevant information had been left out, and 68 per cent rated the results of their interviews "highly satisfactory". All the scientists interviewed agreed that popular news coverage of science issues was important for the public, but only 84 per cent thought it was important for the scientist.

According to John Wallace, journalism co-ordinator at RMIT, these results suggest that the science reporting is more professional than scientists in general believed, although "scientists may be a little less critical of the news process when they themselves are getting publicity."

Tania Ewing

gloomy economic forecasts for 1991-92.

Sir Aaron Klug, director of the Laboratory of Molecular Biology in Cambridge, one of the MRC's largest research units, says that individual scientists planning new research may be "badly affected" by the funding squeeze, although the majority of current work should not suffer. Klug says the Treasury is partly to blame for the financial problems now facing MRC because of its refusal to recognize a "sophistication factor", which means that the price of equipment and reagents has risen faster than the general rate of inflation.

MRC secretary Dai Rees wrote to all MRC research unit directors on 2 November, to explain the staff and equipment freezes, which he described as "exceptional measures". The MRC was last forced to take similar action during the economic recession of 1980-81. At least £1 million of the projected £3.5 million deficit is due to overspending in MRC research units.

Winterton says that units are given a budget at the start of each financial year, and must spend within it. If units overspend, he says, "in a sense, that's their problem".

Peter Aldhous

NUCLEAR WASTE

Military going underground

Boston

PLANS by the US government to open a permanent disposal site for its military nuclear waste took a step forward earlier this month when the Environmental Protection Agency (EPA) allowed a five-year test period to begin at a federal repository tunnelled into salt beds half a mile below the New Mexico desert. The EPA's decision, coming 15 years after work began on the site, is a belated victory for the Department of Energy (DOE), but several vexing problems remain before the site can open officially.

The EPA's ruling exempts the repository — the Waste Isolation Pilot Plant (WIPP) — from the full requirements of the US hazardous waste laws in order for the DOE to proceed with tests.

During the test period, DOE will monitor the wastes to ensure that they do not migrate out of the repository or generate gases that could become volatile if trapped in the underground salt beds that provide WIPP's geological containment.

EPA issued a variance of its normal

rules, which would have forbidden the transportation of radioactive and mixed hazardous wastes to the New Mexico site from nuclear production facilities around the country. Even so, thorny debates over the DOE's title to the land and the method of safe waste transport must be resolved before operations at WIPP can proceed. Also clouding the facility's future is the likelihood of legal action against the facility by local residents, state authorities and national environmental groups.

According to EPA officials, the variance requires the waste to be stored in a "readily retrievable" form during the test, and permits the handling of "no wastes for purposes other than testing". In addition, EPA will limit the amount of waste shipped in the five-year test period to 8,500 drums of material — less than one per cent of the repository's total planned capacity.

WIPP is intended to store transuranic waste, which consists of a wide variety of plutonium-contaminated materials from the production of nuclear weapons, but which does not include high-level nuclear wastes such as the spent fuel rods from nuclear reactors. The transuranic waste, ranging from waste water and sludge to soiled equipment and clothing, is currently piling up in drums in temporary storage facilities at ten DOE facilities around the country. Many of the drums also contain hazardous chemical wastes mixed in with the radioactive materials, a factor that allowed the EPA to exercise some regulatory jurisdiction over the programme in the first place.

DOE officials are relieved to see the WIPP site moving forward.

Late last year, DOE Secretary James D. Watkins described the opening of the WIPP facility as his "top priority". EPA hydrologist Reed Rosnick emphasized, however, that the variance is limited solely to the test phase. "The EPA decision will not alleviate the problems of waste storage at DOE because of the small amount of waste allowed", he said. In addition, Rosnick stressed, if during the test phase it were determined that the WIPP facility cannot meet the required safety criteria, "all bets would be off" as to the facility's future.

The facility has been plagued by unexpected problems since work began in 1975. Worries over excessive water seepage have now been allayed, as have concerns that cracks in the salt formation rendered the tunnels less stable than expected. But the latest worry is that accumulated gas pressure from the decaying waste could far exceed what the surrounding rock can bear, and possibly cause the waste to explode.

On top of these technical concerns, a

variety of administrative and political hurdles remain, not the least of which is that the DOE must acquire title to the land from a branch of the Department of the Interior. Congress failed in its last session to authorize the land transfer, and the DOE is now expected to try to gain title through an inter-agency agreement. But such an arrangement could be open to legal challenge. Meanwhile, a collection of environmental groups, including the Natural Resources Defense Council, is considering legal action on the grounds that the EPA is offering a variance to the US government that would not be given to a private firm.

Federal regulations for the facility also require that the government considers the issue of "human intrusion" which could breach the planned environmental safeguards and potentially prove hazardous to future generations (See *Nature* 344: 576 1990). The issue emerged most recently this spring at a meeting of the panel of the National Academy of Sciences that was overseeing the progress of the WIPP site. So far there has been no agreement on how to mark and protect the WIPP facility against intruders for the next 10,000 years.

Seth Shulman

EUROTAXONOMY

Carrots to lose their roots in fruity future

London

BRUSSELS-based bureaucracy has struck at the carrot, a root tuber that from 1 January 1991 is to be declared a fruit — at least for

Mary Evans

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Rooti fruitti

jam making. The European Communities (EC) Jam Directive insists that jam must only be made from fruit. Jam made from carrots, however, is popular in Portugal. In a concession to a member nation that reveals a certain diplomatic ignorance of botany, EC officials changed the status of the carrot so the Portuguese could still make carrot jam and sell it as such.

Cooks will now ponder the implications of the new move. If the carrot is now a fruit, then ginger — another root used to make jam — must be a fruit as well. Carrots will always be vegetables, regardless of EC directives. But if they are sold as fruits, wherefore the marrow and the tomato, fruits sold as vegetables that are nevertheless used to make jam? Henry Gee

BIOTECHNOLOGY

Bush veto fuels orphan drug act uncertainties

Washington

PRESIDENT Bush last week unexpectedly announced that he would not sign a bill that would prevent companies from reaping windfall profits under the provisions of the orphan drug act, a law that grants commercial and legal advantages to companies that develop drugs for rare diseases (see *Nature* 348, 101; 1990). The bill, which was passed by Congress last month, will not now become law.

Objections raised by Health and Human Services Secretary Louis Sullivan last June — that any change in the law would be a disincentive to investment in the development of orphan drugs — may have influenced the decision taken by President Bush.

Industrial Biotechnology Association president Richard Godown, who was for similar reasons initially opposed to any change in the law, had nevertheless urged the President to sign the bill because he felt that "a genuine compromise had been reached". Godown is now concerned that the biotechnology companies will have to suffer through several more years of additional wrangling over the orphan drug act, which will in turn hinder companies' ability to raise capital and raise the spectre of more restrictive legislation emerging from the next session of Congress.

Diane Gershon

MITI makes its move

Arita

WHILE the western press has noted at length Japan's efforts to develop practical high-temperature ceramic oxide superconductors, less attention has been paid to the even greater effort Japanese industry and government are pouring into the development of other advanced ceramics, notably ceramic/carbon fibre composites that will be strong and light enough to make the outer skin of space planes and hypersonic aircraft. In a talk on composite ceramics at an international conference last week in Arita, on Japan's island of Kyushu, R. W. Davidge of the UK Harwell Laboratory said that "of the 5,000 copies of my book sold worldwide, half have been sold in Japan." Last year, the Ministry for International Trade and Industry (MITI) moved into the area with a new project under its programme of 'research and development of basic technologies for future industries' (jiseidai)¹, and the materials are among the latest subjects for study at MITI's laboratories in Nagoya and Kyushu.

These new ceramics, hard and heat-resistant, have a huge range of applications, from car parts to false teeth. Hundreds of Japanese companies are pushing their development, with activity particularly concentrated in the Nagoya area, the traditional centre of Japan's pottery industry and the home ground of Toyota.

In 1985, over 300 Japanese companies donated a total of ¥5,900 million (\$45 million) to establish the Japan Fine Ceramics Center (JFCC) in Nagoya. With backing from MITI and more money from the local governments of Nagoya, and Aichi and Gifu prefectures, JFCC has become a state-of-the-art centre for ceramics research, performing commissioned research to a value of ¥700 million (\$5.4 million) a year.

Automobile manufacturers are particularly keen to develop lighter more fuel-efficient cars using ceramic materials. Ceramic turbochargers are already used in some Japanese-made cars, but mass production of ceramic engine components has been problematic because of poor yield and disparities in strength. And ceramics cost ¥5,000 to ¥10,000 a kilogram compared with a few hundred yen a kilogram for steel. One of the key aims of JFCC is to help bring down these costs by developing standard testing, inspection and evaluation techniques.

In addition to the JFCC, MITI established several other jiseidai projects in the 1980s to support ceramic research, such as the 'high performance ceramics' programme, focusing on the development of silicon nitride and silicon carbide ceramics for use as high-temperature components in gas turbine engines. Japan, with 50 per cent of worldwide sales,

already dominates the carbon fibre production market, and has strong capability in the development of new materials — both PAN (polyacrylonitrile) and pitch-based carbon fibre were invented in Japan. Yasuhiro Yamada, of MITI's Government Industrial Research Institute in Tosu Kyushu, says that Japan started this kind of research late compared with the United States and Europe. But MITI has organized a research consortium of 12 companies, including Nissan, Nippon Oil Company, and Toa Nenryo Kogyo, an oil refiner that is pushing development of pitch-based carbon fibre, to develop composite ceramics. And given Japan's tremendous strengths in both ceramic and carbon fibre research it will not be long before Japanese companies catch up with the West.

David Swinbanks

COMPUTER RESEARCH

Artificial postmen

Washington

NEXT time you scrawl an address on an envelope, consider what you may be asking an automated mail-sorting machine to do. After satisfying itself that your scribble is indeed an address, it scans for something resembling a postal or zip code. On the rare occasions when it finds one, it then tries to determine a delivery route by deciphering as much of the street address as it can. Needless to say, failure is common; only 4 per cent of handwritten mail (which accounts for some 15 per cent of the US mailstream) can now be machine-read by the US postal service.

Technology now being developed is better — it can read some three-quarters of the handwritten letters it sees — but takes two to three minutes to scrutinize each piece.

Dissatisfied with progress in machine reading after some 15 years of research, the US postal service last week decided to try something new. It announced plans to set up its first "research centre of excellence", a permanent computer vision laboratory at the State University of New York, Buffalo, where researchers have been working on the problem for several years. With initial funding between \$2 and \$3 million per year, the centre is aiming at artificial-intelligence and neural-network-based machines that can read 12 or 13 pieces of handwritten mail per second, says associate director Jonathan Hull. Working prototypes could be ready in as little as a year or two, he says. Among the top research priorities will be the use of "contextual analysis" to decipher a problem character or word based on those around it, and to compensate for missing or incorrect postal codes by making a sensible guess based on other address elements.

Christopher Anderson

Punishment for success

Arita

ALTHOUGH MITI is pouring support into ceramic research, a MITI-backed researcher who made a breakthrough in the development of ceramics earlier this year, found that the system had a very odd way of rewarding excellence.

When Fumihiko Wakai, senior researcher of the ceramic science department of MITI's Government Industrial Research Laboratory in Nagoya (GIRIN) announced the development of the world's first superplastic crystal composite ceramic in March (*Nature*, **344**, 421–423; 1990) he was flooded with enquiries from overseas, particularly from the US defence establishment, according to Sakae Tanemura, senior officer of research planning at GIRIN.

But Wakai complains that although he has been supported with funds of ¥20 million (\$154,000) a year for the past three years his research budget has now been slashed to ¥400,000 (\$3,000) a year. He says that his case is not exceptional and that it reveals defects in the system of research grant allotment in MITI's laboratories.

Wakai says his funding was cut because

his department has received a lot of financial support in the past few years, and the "democratic" system in MITI's institutes now dictates that more money must be funnelled into the other five departments.

In obtaining research funds the chief hurdle to overcome, according to Wakai, is the internal screening of research proposals in the institute before they are submitted to MITI's Agency of Industrial Science and Technology (AIST); for standard research grants researchers cannot apply directly to AIST for support. Wakai is now looking into the possibility of applying for funds through a special interministry programme run by the Science and Technology Agency (STA). Applications for the STA programme do not have to be internally screened in the institute but competition for the limited number of grants is severe.

Wakai contrasts AIST's system with that of the US National Science Foundation, to which researchers apply directly. Grant proposals are reviewed by outside reviewers, sometimes from overseas, a system that Wakai hopes will one day be adopted in Japan.

David Swinbanks

Two into one goes slowly

Munich

LIKE everything else in what was East Germany, the universities were undermined by the need to deal with Communist rule: at least outward accommodation to the Communist system, sometimes escalating to include the betrayal of colleagues to the secret police, was almost always required for even rudimentary academic success. As the leaders of the West German university system, which enjoys remarkable freedom from political interference, take over the running of these decayed institutions, the task of sorting out what and who in the east deserve to be retained is proving difficult and divisive.

Universities enjoy an autonomy unique among public institutions in Germany. Although they are run by the state, universities benefit from a tradition of academic freedom bolstered by post-1968 reforms giving faculties and administrators, not the state, sweeping powers over hiring.

The most politically active university organization in Germany, the newly christened "University Rectors' Conference" (HRK) — previously the West German Rectors' Conference — has included its recommendations for reform in a routine list of demands for more money. In uncharacteristically soft language, it called for the establishment of committees within eastern universities including "significant numbers of external scientists" to examine "structural and personnel

questions." Explaining the delicate approach, HRK president Hans-Uwe Erichsen says that HRK "wants to avoid any sort of politically motivated inspection". Although one may doubt the capability of the universities to purify themselves, he says, it is enough to include foreign researchers in the committees to ensure a fair evaluation. He admits that part of the reason for the careful approach is the fear of endangering university autonomy in the West. "We don't want to set a precedent," he explains, by allowing politicians to dictate hiring policy to university administrators. "What have been called 'partial and temporary' reductions in university autonomy in the past have often not been temporary at all." But in recommendations to be issued this week, the other political heavyweight in the university arena in Germany, the Wissenschaftsrat or science council, is expected to take a different line. According to chairman Dieter Simon, the Wissenschaftsrat will call for the temporary establishment of scientific committees to make binding recommendations to the governments of the eastern *Länder* (states) on how to reform universities.

These committees would be extremely powerful, having the ability to augment present university faculties or recommend entirely new structures. To ensure that old-boy networks are dismantled, the plan calls for each committee to include members exclusively from western Germany, foreign countries, and other eastern

Länder but not from local universities.

Simon expects the plan to be controversial. "Professors at these universities will be understandably critical," he admits, "because they were muzzled before by the state. To them, this just looks like a different style [of repression]." But he argues that the loss of autonomy is the price that must be paid to improve both the scientific quality and the moral standing of the universities.

Simon stresses that the plan can work only if the federal and *Länder* governments provide more money to attract people to new posts in eastern German universities, a demand supported by HRK. Professors in the east still earn as little as one-half to one-fifth of the salaries of their counterparts in the west.

The downside to the plan, Simon recognizes, is that it does nothing to protect those who are politically untainted but whose talents are not far above average. For them, Simon admits unhappily, there is no solution as yet.

The problem of reforming the universities in the east is so daunting that many western leaders, including Erichsen, cling to statistics that show that up to two-thirds of the professors at universities in eastern Germany will retire in the next ten years, ensuring a certain degree of self-purification. But Simon criticizes this optimism, saying that there are plenty of younger people who are badly compromised. Precedents set in West Germany indicate that a faculty member forced unwillingly out of a position has a good chance of using legal action to reverse the decision or at least obtain a substantial settlement, and here again Erichsen is reluctant to introduce any changes in the rules for fear of setting an uncomfortable precedent.

Former East Germans active in the reform process say the problems will persist for decades. "A committee passing out 'certificates of purity' will certainly not solve the problem," says Guntolf Herzberg, a philosophy lecturer at the Free University of Berlin who was expelled from East Germany in 1985. According to Herzberg, the committee's checks will inevitably be superficial because real research into people's backgrounds under the former Communist rule requires "almost detective-style work" based on files that in many cases have disappeared. The problem, he says, parallels that of the larger society: there were an estimated 300,000 people working for the Stasi (secret police), and "you can't lock them all up". Herzberg also fears that harsh legal measures to "filter out the guilty ones" would have severe consequences for a young democracy. Therefore, he says, it is better to give people the benefit of the doubt even when there are strong suspicions against them.

Steven Dickman

A tale of two countries: divisions to be bridged

Munich

GERMANY is now a single country with two university systems, which differ in the following ways:

Class size. Whereas lectures and laboratory courses in the west are hopelessly overcrowded (especially in popular fields like biology), students in the east attend class under nearly idyllic circumstances. Student-faculty ratios are around 5-to-1 or even less in the east; in popular fields in the west they may be as high as 30-to-1.

Who may study. During the 1960s, West Germany relaxed entrance standards to its universities and today 25 per cent or more of those in a certain age group attend university. East Germany selected its students much more strictly at an early age, with the result that only 13 per cent of students of a given age enter university.

Content of courses. Although change is coming, many courses in the east are still taught using textbooks or curricula developed before the Communist government collapsed. Especially in the humanities and social sciences, professors have put together a patchwork of readings and assignments that only "after a fashion" live up to the quality and consistency of courses at univer-

sities in the west, according to Dieter Simon, chairman of the Wissenschaftsrat.

Access to research facilities. Despite complaints from many western German students and professors, they are blessed with some of the most modern and well-maintained university research facilities in the world. Not so in the east, where years of government favouritism toward institutes of the Academy of Sciences (now disbanded) drained the universities of resources, not to mention the best researchers.

Attitudes. Universities in the east were barely involved in the 'revolution' that led to the fall of the Communist regime, with students taking a conservative view in order to protect their state benefits. "It is really disturbing," says Guntolf Herzberg, a philosophy lecturer expelled in 1985 from the Free University of Berlin; "there have been too few cases where students have risen up even to question their professors' attitudes, let alone to throw them out." "The real difference is in the way people think," says Simon; where the West was americanized, the east Germans "remained much more German than we did — respectful of authority and distant from those around them — which is very unfortunate." **Steven Dickman**

Dissent hits climate accord

Geneva

THE 137 countries represented at last week's World Climate Conference left Geneva having agreed to begin negotiating towards an international convention on climate change problems, with the aim of signing an agreement at the United Nations Conference on Environment and Development to be held in Rio de Janeiro in 1992. But even at this preliminary stage there is disharmony: a newly formed coalition of the small island states, perceiving an imminent threat from any rise in sea level, made known its disappointment with the ministers' final conference declaration.

Hours before the ministerial declaration was adopted the Alliance of Small Island States (AOSIS) came into being. Delegates from the AOSIS nations agreed to adopt the final declaration in Geneva, but only after expressing support for an alternative version that contained stronger wording on the threat to low-lying islands and the need for precautionary measures to combat global warming. Lincoln Myers, environment minister of Trinidad and Tobago, said he was angry at being treated as "a codicil" to the meeting, and said AOSIS would "be extremely firm in the negotiations to come", pushing for limits on carbon dioxide emissions from the industrialized countries.

Although the co-ordinated effort will strengthen the hand of the island states, it is still expected that the greatest role in the negotiations (which will begin in Washington next February) will be played by the three nations most strongly opposed to carbon dioxide emission controls (the United States, the Soviet Union and Saudi Arabia) and those industrialized countries which have already adopted targets to stabilize or reduce their emissions. A binding protocol to limit carbon dioxide emissions still seems unlikely before 1992. John Knauss, under secretary in charge of the National Oceanic and Atmospheric Administration, and head of the US delegation in Geneva, said "we ought to reach an agreement on a framework convention first and then discuss protocols".

Another difficult issue is the transfer of technology and funds to poorer countries to allow them to develop industrially while limiting greenhouse gas emissions.

As expected, the conference declaration falls short of urging industrialized countries to set targets that would stabilize carbon dioxide emissions by the turn of the century. But the declaration mentions by name those countries which have set such targets, and the US delegation failed to gain recognition for US efforts to limit total greenhouse gas emissions. In diplomatic language, this indicates growing

"irritation" with the US position, according to a member of the British delegation.

Brice Lalonde, the French environment minister, explained that the European Communities countries had wanted to

avoid isolating the United States before the start of convention negotiations, and so had not fought hard for stronger statements on carbon dioxide. But the United States must soon begin to alter a way of life "founded on the low cost of energy", he said, and increase energy prices to reduce carbon dioxide emissions.

Peter Aldhous

Iron solution no solution

Washington

AN ambitious proposal to remove excess carbon dioxide from the atmosphere by using iron fertilizer to stimulate the growth of gigantic algal blooms was dealt a blow earlier this month when a group of experts determined that the technique would be hugely expensive, potentially dangerous, and would only reduce total carbon levels by a small amount.

The proposal surfaced earlier this year as a possible last-ditch response to a runaway increase in carbon dioxide levels and the accompanying rise in global temperatures caused by the greenhouse effect. In a widely reported article, John Martin, of the Moss Landing Marine Laboratory (see *Nature* 345, 374; 1990), hypothesized that the growth of algae in the ocean around Antarctica is limited by a shortage of iron. Simply adding iron to the water could prompt massive algal growth, Martin wrote. He calculated that addition of 300,000 tonnes of iron would be sufficient to remove 2,000 million tonnes of atmospheric carbon, which is equivalent to about a third of the yearly output of the world.

But at a recent conference in Irvine, California, with sponsorship from the US National Research Council (NRC), scientists presented new data that suggest more iron will be needed and less carbon can be extracted. The best guess is now that between one and five million tonnes of pure iron would be required to stimulate the growth of algae capable of soaking up 1,000 million tonnes of carbon. That figure assumes that iron — rather than some other nutrient — is indeed the limiting factor in antarctic algae growth, something that has not so far been conclusively established. Those calculations represent an increase of some three to fifteen times the amount of iron level specified that Martin specified in his original calculations, with only half as much carbon likely to be extracted.

Compared with the current world output of some five to six million tonnes of carbon, savings of 20 per cent may not be worth the trouble, says Karl Banse, a biologist at the University of Washington who served on one of the panels of the conference. "Maybe we would be better off investing our money somewhere else",

Banse suggests.

Neglecting even the environmental and practical uncertainties of such a project, economics remain a major question mark, says Thomas O'Brien, a technology assessment expert for US chemical company. DuPont Crude calculations, he suggests, show that the iron itself — in the form of a 30 per cent solution of liquid ferrous chloride — would cost between \$150 and \$200 a tonne; the equivalent of five million tonnes of pure iron could cost as much as \$3,000 million.

Although ferrous chloride happens to be a natural waste product of other DuPont production processes, such a massive demand would quickly outstrip the levels that result from routine production, forcing the company to make it for its own sake — at even higher prices. But even those costs, he says, would be "insignificant" compared with the expense of transporting the caustic chemical to the Antarctic ocean in hundreds of rubber-lined barges.

Troublesome, too, are the questions of the environmental impact of sudden algae blooms. Biologist Gustav Paffenhöfer, of the Skidaway Institute of Oceanography, says that for krill, a small crustacean that plays a vital part in the antarctic food chain, huge algae fields "could be the best thing that ever happened — or the worst". Krill feed on algae, but they also lay their eggs in the water below the surface. Both algae and eggs slowly sink, the eggs eventually hatching while the algae deteriorate. If algal decay creates oxygen-depleted strata at the same depth as that at which the eggs hatch, entire generations of young krill could be killed. Likewise, tunicata, jellyfish-like organisms that eat both small krill and algae, could "take over", unbalancing the food chain, he says.

Nevertheless, the conference participants agreed that the 'iron solution' — even if it is never the answer to the threat of global warming — is worth further study. They suggest two years of preliminary laboratory and field studies, followed by a trial iron-fertilization experiment covering some 400 square kilometres of ocean. NRC is considering a formal study and report, which could be completed late next year.

Christopher Anderson

Grass-roots not so green

Washington & San Francisco

AFTER a summer of growing environmental activism, voters across the United States rejected virtually every 'green' measure on the state ballots in last week's national elections. In the six states where environmentalists had concentrated their efforts, voting went by as much as two to one against referendums that would have stiffened regulations of emission, logging, recycling, pesticides, the nuclear power industry and waterways. All of the most sweeping initiatives — notably California's 'Big Green' proposal and New York's \$2,000 million environmental bond issue — were defeated.

Supporters of the referendums attribute the losses to massive opposition campaigns sponsored by industry and other opponents.

In Oregon, the oil and chemical industry outspent environmentalists eight to one over a recycling initiative, eventually drowning the measure in a \$2.5 million television advertising campaign. The less affluent supporters of the referendum were forced to turn to the federal 'fairness doctrine', which requires broadcast media to provide "reasonable opportunity for response" to political advertising. But that typically meant one hour of free pro-initiative air time for every four hours bought by opponents", says Jon Stubenvoll, press secretary for the consumer recycling coalition that sponsored the referendum.

Industry-sponsored advertising also showed a shift towards hard-line tactics. The Oregon recycling initiative had been ahead in the polls, Stubenvoll says, until opponents aired a commercial that implied that restrictions on plastic food wrap could result in outbreaks of botulism, hepatitis and salmonella. "That's hard to fight. Fear is an incredible tool", says David Hamilton, national field director for the US Public Interest Research Group, which helped sponsor some of the initiatives.

Despite the defeat of specific proposals, environmental groups did well in backing the campaigns of pro-environment politicians.

In Congress, pro-environmental candidates picked up 14 seats — 13 in the House of Representatives and one in the Senate — according to Reid Wilson, political director of the Sierra Club. He also calculates a net gain of five state governorships on environmental issues.

With stronger environmental representation at the congressional and state levels, grassroots initiatives of the sort that failed last week may become less necessary, Wilson says. "Initiatives are started because the voters are frustrated by inaction on the part of their elected

officials. The more good officials we have, the fewer initiatives we'll need." In California, both major environmental propositions — Big Green (see *Nature* 347, 323; 27 September 1990) and a timber measure known as Forest Forever — lost last week, but so did their industry-sponsored counter-measures, one proposing compromise regulations on pesticides and another creating timber bonds.

The pattern that emerged across the United States is that virtually "every proposition that called for significant outlays of money was defeated", according to Ralph Cavanagh, energy program

director at the Natural Resources Defense Council.

Cavanagh says Big Green was hurt most by opponents who framed the measure as one which would lead to higher food and energy costs.

"If the elections had been held before the Middle East crisis and the budget battles in Congress last month, I think that many of these initiatives would have won", Wilson says. But given that economic uncertainty is likely to be a feature on the political landscape for some time, supporting pro-environment politicians rather than ballot referendums may prove to be a stronger strategy in the future, he says.

Christopher Anderson & Elizabeth Schaefer

GREENHOUSE GASES

Swiss cut it to the bone

Munich

In a fitting coda to the World Climate Conference in Geneva, the host government announced on 31 October that it intends to pursue what will be one of the world's strictest policies. Swiss environment Minister Flavio Cotti introduced a government proposal for a stiff carbon dioxide tax and other measures aimed at reducing CO₂ emissions by around 2.5 per cent before the year 2000.

Although Switzerland contributes a paltry 0.2 per cent to world CO₂ output, the government hopes to set an example to its neighbours. The Swiss plan, which has to be approved by the parliament in mid-1991 in order to take effect, aims to reduce consumption of fossil fuels by imposing huge tax increases. Coal will be taxed at 42 to 105 per cent of the market price, petrol at 15 per cent, heating oil at 23 per cent and natural gas at 20 per cent. And Cotti said, the tax would probably have to be doubled in 1995 to achieve the intended goal.

The CO₂ tax, if approved, would be

about double a tax being considered in Germany and several times higher than a tax already in effect in the Netherlands. The only other European country taking a comparable stand, according to a Swiss government official, is Sweden. The official said the tax proposal was received "quite positively" by the parliament and the public, which is also enthusiastic about reducing emissions of nitrogen oxides that are destroying Swiss forests. He expects the proposal to survive in some form despite probable industry opposition.

Ironically enough, it is possible that the parliament will reduce the tax not because it is too high for consumers, but rather because it would bring in too much money — an estimated 1,900 million Swiss Francs (about \$1,500 million) a year — to the already overflowing government coffers. The government surpluses that have become the norm in Switzerland are not allowed to go above certain levels, so that every new tax must be accompanied by a reduction in an existing tax.

Steven Dickman

ANIMAL RIGHTS

Dolphin excused military service

Boston

WITH the help of public demonstrations, front page newspaper stories, local television reports and a lawsuit, animal rights

dolphin had been scheduled for transferral to a classified programme in the U.S. Navy which reportedly trains marine mammals for military work. News of the transfer ignited strong popular sentiment, with a former Navy dolphin trainer weighing in on Rainbow's behalf and attacking the Navy's programme.

According to workers at the aquarium, Rainbow, an 11-year-old male, had been put on the transfer list because he showed increasingly aggressive tendencies towards his companions. Aquarium officials have now promised that Rainbow will not be enlisted in the Navy, although what will be done remains unclear. Seth Shulman

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Dolphin draft-up

activists have staved off the prospect of military service for Rainbow, a dolphin at the New England aquarium in Boston. The

The buck stops with government

SIR—The recent debate on the closure of several major research groups by the MRC has mainly focused on the way in which the council has taken these decisions. I have recently been on one of the MRC research boards, and from 1980 to 1990 I chaired one of the MRC committees that rank applications for project grants. From that experience, I believe that much of the criticism should instead have been aimed at the government.

Although I have sometimes disagreed with the priorities adopted by the recently formed MRC strategy committee, the government is responsible for putting the MRC in a position that forces it to make awful choices and to fund Britain's excellent medical research community inadequately. As a result, we have lost excellent science not only in a few high-profile MRC units, but also in laboratories deserving support through long-term programme grants or short-term project grants.

In many ways, the most depressing of these changes — both for the MRC committee members and their clients — is the striking decline in the funding of project grants. Although it is easy to focus public criticism on research council decisions to close major research enterprises, the recent across-the-board decimation of excellent project grant proposals has been a much less public 'death by a thousand cuts'.

The loss of project grants may in the long term have disastrous consequences for the whole UK research enterprise. Two things tend to be lost. First, many new and bright young scientists setting out on research careers — who hold the future of British science in their hands — fall at the first grant-seeking hurdle. Second, funding is disappearing for many of the more senior and still excellent academics who combine first-class research with increasingly heavy teaching and administrative roles in financially strapped universities — but whose research is on a scale too small or in a field too unfashionable for programme grant funding.

A year ago I took up these concerns in a letter to the prime minister. Her first response was to point out that there was really no problem, as MRC "in recent years" had funded about 80 per cent of α -rated project grants — which was true for a couple of years in the mid-1980s — and that new money made available in 1988–89 "should allow [MRC] to sustain or increase this level of support for grants". When I informed her that in the event only 55 per cent of the 1988–89 α -rated grants were funded, a second letter from Downing Street told me that this fall "cannot be attributed to government". The MRC had been given more money and "it is for the MRC to deter-

mine its own funding priorities".

During 1989–90, both the absolute number of project grant applications and the number of grants rated by the grants committees increased substantially over 1988–89, but the number funded was essentially unchanged. The funding rate fell to 47 per cent of α -rated proposals, little more than half of the 80 per cent that in September 1989 was deemed a desirable target by the prime minister.

The real problem appears to be that the UK government still believes that it can get first-rate research without first-rate investment. And the research councils, who act as the government's agents in distributing research funds, have failed to emphasize just how much of the good science they should be supporting is lost for lack of funds: too often they greet modest increases in funding with statements of public gratitude rather than with reminders of the problems that remain.

R. H. MICHELL

*School of Biochemistry,
The University of Birmingham,
Edgbaston, Birmingham B15 2TT, UK*

SIR—In the news item 'Storm over funding policy' (*Nature* 348, 6; 1 November 1990) much prominence is given to criticisms of recent closures and cuts made by the MRC. There also was some suggestion that the strategy committee of the MRC council has acted in a way contrary to the normal practice of basing funding decisions on peer review. I would like to put forward a different view based on my experience as a past member of the MRC council and a member of the strategy committee from its inception until my retirement from the council earlier this year.

First, the MRC faces financial problems not of its own making. Quite simply, if there continues to be an annual gap between the increases in cost and the provision for salaries in the grant-in-aid, as there is in 1990–91, staff will need to be reduced by 6 per cent every year unless savings can be made elsewhere. Indeed, if this gap extends to all research costs, as it does, then the MRC will be forced to curtail more and more of its activities. In times of financial stress it is short-term support, like project grants, that can be most easily cut, and so the science that will suffer most is that carried out in universities and medical school laboratories. This is unacceptable because it destroys the research base for graduate student training in the universities with disastrous longer term effects.

The MRC council set up the strategy committee to deal with these problems and to avoid what has happened in the past when urgent financial decisions were made administratively without much

regard for scientific policy. The committee includes the chairman of the boards and two to three of the independent scientific members of the council. It is definitely not a group that can act independently of the boards; indeed, its structure ensures that decisions are based on proper scientific review. And, of course, all its decisions are subject to approval by the whole council. None of the recent decisions has been taken peremptorily; all are the outcome of the development of long-standing policy.

The council has not in any way abrogated the processes of peer review. Experts in a particular field can make assessments only within that field and different judgements are required for comparing work that ranges over different fields. As many prospective authors to your journal well know, passing technical referees even with flying colours does not guarantee appearance. Additional appraisals in wider fields are made by the council's boards, and the strategy committee was set up to pursue the same process at the most general level and to ensure that funding decisions are still subject to scientific policy and not made by accountants.

For all its much vaunted success peer review has a dark side. When money is flowing, it is marvellous because all it has to do is to exclude the extremes of lunacy and banality. As funds contract, there is not only greater and greater selection of the conservative and the proven successful, but this is accompanied by a growing tendency of applicants to cast their proposals in the same conservative mould. This threatens innovation, which is always risky, and affirmative action is needed by the MRC and other institutions to encourage new research and, if necessary, to override expert opinion. If, 40 years ago, the MRC had relied on the judgement of contemporary experts they would never have supported the infant science of molecular biology.

SYDNEY BRENNER

*MRC Molecular Genetics Unit,
Hills Road, Cambridge CB2 9QH, UK*

Verbal abuse?

SIR—In a recent article (*Nature* 347, 225; 1990), John Maddox protested rather trenchantly (albeit parenthetically) to using the verb form of the noun "intuition". I, for one, intuit (from the Latin *intuitus*, past participle of the verb *intueri*) no good reason to prohibit (from the Latin *prohibitus*, past participle of the verb *prohibere*) such use, and I pray he does not edit (from the Latin *editus*, past participle of the verb *edere*) the life out of this riposte.

GEORGE A. DIAMOND

*Cedars-Sinai Medical Center,
Los Angeles,
California 90048-1869, USA*

Does the literature deserve the name?

The general opinion that the scientific literature does not deserve the name is widely accepted, but those chiefly responsible, the authors of research reports, seem unable to change their ways.

MOST scientific writing is narrative, essentially the stuff with which Joseph Conrad won fame and fortune a century ago. A person wishing to describe the outcome of a series of experiments, or of observations, and to suggest what they signify, has also to tell a tale. But the task is necessarily simpler than those that Conrad chose to face. Imagination, or too much of it, is neither necessary nor permissible (for obvious reasons). And the authors or research reports are not required to strive for dramatic endings. Indeed, they would be out of place. So why is the narrative or research so often laboured and, more mystifying and even more important, so difficult to understand?

A frequent defence is that the scientific literature is so stylized that authors' narrative talents are inhibited by the straitjackets into which they are forced, but this can hardly be the case. It is true that some journals demand that research should be described under various headings, METHODS, METHODS AND MATERIALS, CONCLUSIONS and — of fine distinction — DISCUSSION, but that should make the task of telling a simple tale easier, not more difficult. If the structure of the narrative is predetermined, should not its elements be more easily told in simple language?

What seems to happen, instead, is that a pre-determined format is taken as an instruction (or an excuse) for replacing narrative prose with a listing. Sentences without verbs become commonplace. Abbreviations, not simply for commonly used reagents, but for materials first defined elsewhere in the same article, are liberally sprinkled among the intelligible words. The assumption seems to be that people will not wish to read these stylized sections of a research report as if they were pieces of the narrative, but will refer to them only in the quest for information on, say, the conduct of an experiment.

For what it is worth, many of the legends accompanying figures in this journal are not very different, although we take some care to provide each sentence with a verb if the author has overlooked the need for one. Omissions that cannot as easily be remedied are of the phrases that would help readers quickly grasp why some experimental procedure has been followed. All too often, for example, people will say merely that they have incubated some system Q "in the

presence" of a radioactive material without adding some qualifying phrase such as Q "to label the product".

"But people in the field will know what is going on, and the others will not care", is the customary defence. But can that be certain? While those familiar with experiments of the kind described may indeed be able to understand what is going on, will not even they benefit by being able to read more quickly from explicit pointers? And might there be some merit in capturing the interest of people working in quite different fields? With that said, of course, the now-familiar phrase "Quantification was achieved by . . ." deserves to be banned.

The cramping consequences of the stylized nature of some communications are probably exaggerated by those who mistake a listing for literature, but that is not the reason most commonly given to explain the obscurity of the literature, nor should it carry much weight with hapless readers that authors often say that space is at a premium, and that the demands of journals require of them feats of compression surpassed only by the better poets. The truth is that most editors would willingly give extra space to well-written articles, that the avoidance of repetition by the judicious use of pronouns would usually save more space than is consumed by phrases meant to say what experimental steps are for, and that the most common thief of space is the narrative of material only peripherally relevant to the purpose of an article.

Much the most common defence of the obscurity of the literature is that it is intrinsically obscure, but that is at best only half an excuse. Of course, the literature is by definition full of articles that describe previously unknown phenomena, or which embody data whose significance has not yet been fully appreciated. It is also well-known, among narrative writers, that it is easier to tell a clear tale than one whose purpose is unclear. So should it not be the rule, rather than the exception, that scientific articles are clouded?

Only if one accepts that it is not possible to make clear statements about circumstances that are themselves uncertain. But that is plainly not the case. Even if a person has discovered something truly puzzling — perhaps, to take a classic case, that the bombardment of uranium with neutrons yields short-lived β activity

as well as α activity — there should be no difficulty in saying that "there is at present no explanation of these observations, which are nevertheless easily distinguishable from the background noise and also reproducible". (That is more or less what Fermi said about his experiments with uranium; it remains a puzzle, and a surprise, that he missed the explanation.)

The truth is that badly written papers are most often written by people who are not clear in their own minds what they want to say, and which seem also to derive some of their obscurity from extraneous but important issues that complicate the tale an author has to tell. If, for example, a person is describing experiments that would not have been possible without the help of a third party, perhaps in the provision of a crucial piece of information, an author will tend to beat endlessly about the bush, seeking to make clear to interested readers that his colleague's help had indeed been crucial without at the same time giving readers the impression that his own contribution had been unimportant.

It would be an interesting exercise to see what would happen to the unfolding content of the literature if journals were to follow the policy of publishing only clearly understandable articles. Would that be at once a guarantee of quality and an assurance that authors have nothing to hide?

In reality, this journal offers would-be Conrads opportunities too often declined. There is no set format, but a general understanding that articles will begin with a statement of a problem and a recitation of its antecedents, historical and otherwise, that it will continue with an account of what has been done and that it will end with some discussion of what has emerged — a conclusion ("Water flows uphill"), a qualified conclusion ("Water flows uphill only when the hill is spun rapidly in a centrifuge") or an allusive suggestion that others may find helpful ("It has not escaped our attention that the phenomenon we describe may be relevant to the provision of public water supplies on the surfaces of rotating neutron stars"). As it should, the recipe has a beginning, a middle and an end. Conrad would have approved. When there is so much exciting discovery in all fields of science, it seems a shame that so much of it should be buried because the recipe is not followed.

John Maddox

Gating for the physiologist

Bruce Bean

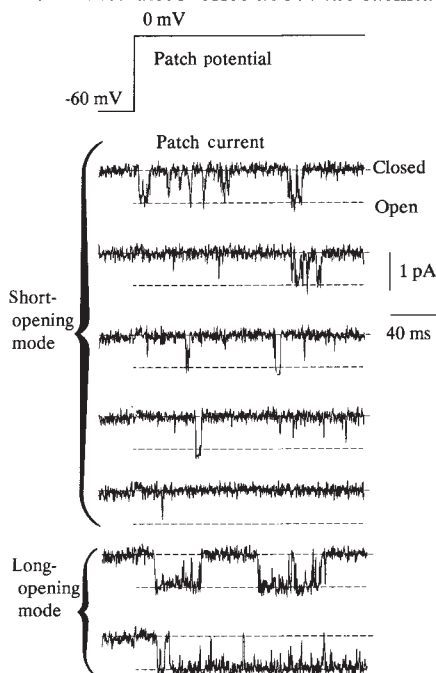
VOLTAGE-dependent calcium channels are present in the membranes of almost all excitable cells. Such channels are closed at normal negative resting membrane potentials, and are 'gated' by membrane depolarization into open, Ca-permeable states, which are usually short-lived. Since the invention of the patch-clamp technique¹, a variety of unusual gating behaviours has been revealed, ranging from periodic episodes of uncommonly long channel openings (see figure) to long-lasting after-effects of brief voltage pulses. Although they provide interesting puzzles for ion-channel biophysicists, most of these peculiar phenomena have been of uncertain relevance in helping us to understand the physiological functions of Ca channels. Now, on page 239 of this issue², Artalejo *et al.* show that one such phenomenon is intimately related to a mechanism by which gating is controlled by neurotransmitter stimulation of the cell. Because transmitter modulation of Ca channels is an important element of physiological control pathways in the cardiovascular, endocrine and nervous systems, their findings will transform the details of Ca channel gating from biophysical arcana into a subject of general physiological interest.

The essential functional property of all voltage-dependent channels is that they are opened by very small changes in the membrane potential. Under normal circumstances, the voltage dependence of Ca channel gating is very steep: channels are completely closed at normal resting potentials of -70 mV or so but are maximally opened if the membrane is depolarized to $+20$ mV or so. This is just the range that is spanned by an action potential in an excitable cell. Most Ca channels open very rapidly, within a millisecond or two, when the membrane is depolarized, and close equally quickly when the membrane is repolarized. These properties permit the depolarization produced by an action potential to induce a quickly developing influx of Ca that can be terminated equally quickly. If depolarization is sustained, a channel normally oscillates rapidly between closed and open states, with openings lasting about a millisecond or two on average (for example the short-opening mode shown in the figure).

In the first study of vertebrate Ca channels by the patch-clamp technique, Fenwick, Marty and Neher³ found an unexpected departure from this normal pattern of gating. They used adrenal chromaffin cells, excitable endocrine cells whose secretion of adrenaline is regulated by internal Ca. Membrane depolarizations beyond 0 mV had peculiar after-effects on

the Ca current in these cells: channels closed quickly following such depolarizations; but, on subsequent depolarization, the Ca current increased. This 'facilitation' outlasted the normal millisecond to millisecond gating of the channel.

Artalejo and colleagues have now made two further discoveries about the facilitation



Spontaneous conversion of calcium channel activity from normal short-opening mode to long-opening mode⁵ in a patch of membrane on a bullfrog ventricular cardiac muscle cell. The channel is closed at -60 mV and is induced to open by depolarizing the patch to 0 mV. Current flowing through the channel into the cell (downward direction) is carried by 115 mM Ba^{2+} . No drugs or transmitters were present. Last six records were taken sequentially, with depolarizations two seconds apart. (pA is picoamperes.)

tion of Ca channels in chromaffin cells^{2,4}. The first is that the facilitated Ca current flows through a completely different set of channels from that carrying the normal current⁴. Facilitation recruits a population of channels (belonging to the 'L-type' class of Ca channels) that, unlike the channels underlying the basal current, are blocked by the dihydropyridine nisoldipine. Evidently, chromaffin cells possess a population of L-type channels that are virtually silent normally but that can be recruited by a large depolarization. The second discovery, even more striking, is that the same population of 'latent' L-type channels can, in the absence of large depolarizations, be induced by dopamine stimulation of the cell². Acting through D_1 receptors, the effects of dopamine seem to be mediated by protein kinase A, because

they can be mimicked by analogues of cyclic AMP and prevented by kinase A inhibitors.

The discoveries of Artalejo *et al.* bear on two recent developments in understanding unusual Ca channel gating in a different kind of cell, the cardiac myocyte. The predominant Ca current in cardiac muscle is carried by L-type Ca channels that can be blocked by dihydropyridine drugs. Normally, the channels show a standard gating pattern, with open times that average a millisecond or so. If the behaviour of a single channel is monitored for several minutes, however, channels occasionally change to another mode of gating characterized by much longer open times⁵ (see figure). Under normal conditions, the frequency of entering this mode of gating is so rare that the importance of the phenomenon — and even its existence — have been doubted. But Pietrobon and Hess⁶ have found that strong depolarizations induce the channel to enter the long-opening mode and that the resulting occupancy of that mode outlasts the depolarization itself. Even more intriguingly, Yue, Herzig and Marban⁷ have demonstrated that stimulation of β -adrenergic receptors on the cardiac myocyte (or application of membrane-permeant cAMP analogues) also greatly increases the frequency with which the channel enters the long-opening mode.

Taken together, the two sets of observations are strikingly reminiscent of those of Artalejo *et al.* on chromaffin cells: a novel form of channel gating can be favoured either by large depolarizations or by cAMP analogues. The analogy goes even further, because in chromaffin cells the L-channel activity induced by dopamine or depolarization is characterized by unusually long openings^{2,8}. In trying to understand how these two different stimuli produce much the same effect on the gating machinery of the channel, an important clue will be the similarity of the induced gating to the long-opening modes produced by the dihydropyridine 'agonist' drug BAY K 8644 (refs 5, 9 and 10), which itself can activate the facilitation current in chromaffin cells⁴.

In both chromaffin cells and cardiac myocytes, it is a dihydropyridine-sensitive L-type channel that is modulated. An

1. Hamill, O.P. *et al.* *Pflügers Arch.* **391**, 85–100 (1981).
2. Artalejo, C.R., Ariano, M.A., Perlman, R.L. & Fox, A.P. *Nature* **348**, 239–242 (1990).
3. Fenwick, E.M., Marty, A. & Neher, E. *J. Physiol., Lond.* **331**, 599–635 (1982).
4. Artalejo, C.R., Dahmer, M.K., Perlman, R.L. & Fox, A.P. *J. Physiol., Lond.* (in the press).
5. Hess, P., Lansman, J.B. & Tsien, R.W. *Nature* **311**, 538–544 (1984).
6. Pietrobon, D. & Hess, P. *Nature* **346**, 651–655 (1990).
7. Yue, D.T., Herzig, S. & Marban, E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 753–757 (1990).
8. Hoshi, T. & Smith, S.J. *J. Neurosci.* **7**, 571–580 (1987).
9. Kokubun, S. & Reuter, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4824–4827 (1984).
10. Brown, A.M., Kunze, D.L. & Yatani, A. *Nature* **311**, 570–572 (1984).

important difference between them is that in cardiac cells the channels contribute a large amount of current (crucial for triggering the normal contraction of the muscle for every heartbeat) by normal gating under unstimulated conditions, whereas in chromaffin cells the current through L-type channels is virtually non-existent until it is induced by strong depolarization or dopamine.

Because L-type calcium channels are widespread, from pancreatic B cells to neurons to fibroblasts, these various

reports are certain to spark tests for similar modulatory mechanisms in other cell types. Analogous pathways might modulate other classes of calcium, sodium or potassium channels. It will be especially interesting to see if this type of control mechanism is included in the large repertoire of modulatory effects at work in the nervous system. □

Bruce Bean is in the Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA.

QUANTUM MECHANICS

Exclusion principle still intact

Tony Sudbery

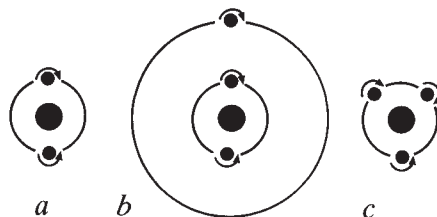
ONE of the greatest successes of modern physics was to establish a basis for explaining all of chemistry in the theory of atomic structure, in which Pauli's exclusion principle is fundamental. Nevertheless, it has recently been queried whether this principle is universally true. On page 224 of this issue¹, Kekez, Ljubičić and Logan report on experimental searches for violations of it, concluding that if the exclusion principle ever is violated, it happens less often than once in 10^{34} times.

In its original form, Pauli's exclusion principle was an adjunct to Bohr's old quantum theory, which stated that an electron in an atom could only move in one of a discrete set of orbits; Pauli added the postulate that there could be only one electron moving in each of the allowed orbits (including the spin direction in the orbit's definition). The different properties of atoms with different numbers of electrons are understood in terms of the different orbits of the electrons (see the figure).

When Bohr's theory gave way to modern quantum mechanics, in which the picturable orbit of an electron in an atom is replaced by the more abstract notion of a 'state' of the electron, the exclusion principle survived as the statement that no two electrons could be in the same state. In this form it applies to all electrons, not only to those in atoms, and also to other material particles like protons and neutrons. It does not, however, apply to photons, the particles representing the quanta of energy in the electromagnetic field; these, on the contrary, tend to cluster together in the same state.

It is now accepted as a general principle that every species of particle is one of these two types: either, like electrons, the particles in the species exclude each other from being in the same state, or, like photons, they tend to congregate in the same state. Particles of the first type, usually regarded as particles of matter, are called 'fermions'; those of the second type, quanta of field energy, are 'bosons'. The theoretical description of the two

types, due to Fermi and Bose respectively, is that the mathematical wave describing the particles undergoes a change of sign if two fermions change places, but is



Pauli exclusion principle: the Bohr-type orbits of electrons in normal atoms of helium (two electrons; a) and lithium (three electrons; b). The different chemical properties of lithium are caused by the more energetic orbit of the third electron. If the uncertainty principle was violated this electron could emit a γ -ray and fall into the lower orbit, forming a non-paulian atom of lithium (c) with two electrons in the same state. This atom would have the chemical properties of helium. Searches have been made, unsuccessfully, both for the γ -rays emitted in processes like this and for atoms of one element showing the chemical properties of another⁸.

unchanged if two bosons do so; this means that fermions are associated with a factor of -1 , bosons with one of $+1$.

This principle is not absolutely fundamental to the quantum-mechanical description of nature; one can imagine a consistent quantum mechanics in which particles are neither fermions nor bosons. This becomes more difficult if the theory is to satisfy the requirements of special relativity and locality (meaning action at a distance is prohibited), which can hold only if all the particles described by the theory are either fermions or bosons. (Interestingly, this is true only if space has three or more dimensions; if space were two-dimensional there could be particles called anyons with properties in between those of fermions and bosons, for which the characteristic factor $+1$ or -1 could be replaced by any number. Such particles might be formed in two-dimensional materials, such as those which manifest super-

conductivity at high absolute temperatures, and there is speculation that anyons might provide the explanation of this phenomenon.)

Nevertheless, in 1987 the Russian physicists A. Yu. Ignatiev and V.A. Kuzmin² described a quantum-mechanical model of particles which almost obeyed the exclusion principle but did not quite do so; there was a small probability that two of their particles might be in the same state. In America, O. W. Greenberg and R. N. Mohapatra³ attempted to develop this model into a three-dimensional theory compatible with relativity and locality. The publication of their work prompted a renewal of experimental searches for violations of the exclusion principle.

According to the general theorems mentioned above, the Greenberg-Mohapatra theory should not be possible, and sure enough it was found to be inconsistent⁴, leading to negative probabilities for some processes. But the experimental searches were not called off on this account. This provides an interesting sidelight on the relationship between experiment and theory in the development of science: a theoretical suggestion was sufficient to prompt the experimental work, but is not held to be necessary for it. Once the idea had occurred that the exclusion principle might not be true, it is worth testing it if one can, whether or not any specific alternative is in view.

The first place to look for violations of the exclusion principle is where it first made its appearance, inside atoms. In an atom with many electrons, some of the electrons are forced into high-energy orbits because all low-energy orbits are already occupied. But if two electrons can sometimes be in the same state, it would seem that an electron in a high-energy orbit will occasionally get an opportunity to lower its energy by joining another electron in a lower orbit, emitting the surplus energy as a γ -ray. However, it is not quite as simple as that. Even if the exclusion principle does not govern all electrons, quantum mechanics will not allow an electron that is obeying the principle to start to disobey it. If we are to see electrons in the desirable inner orbits of atoms being joined by guests in defiance of Pauli's rule, the newcomers must already be lawbreakers and must come from outside the atom. One possible source of such non-paulian electrons might be an electric current flowing around the atom. A recent experiment⁵ by E. Ramberg and G. A. Snow found that fewer than 1 in 10^{26} of the electrons in an electric current in copper were in breach of the exclusion principle.

Very similar considerations apply to the protons and neutrons inside the atomic nucleus. Here the source of the non-paulian particle could be an interaction between a proton in the nucleus and an electron outside it, which could fall into

the nucleus and convert the proton into a neutron. The neutron, being a new particle, might be created in a non-paulian state; it could then join another neutron in a lower-energy state inside the nucleus, emitting a γ -ray. The new test by Kekez *et al.* is based on an examination of the existing data from the underground Liquid Scintillation Detector below Mont Blanc, in which 90 tons of liquid hydrocarbon have been monitored since 1984 for evidence of nuclear events. They find that the probability of a non-paulian event is less than 1 in 10^{34} .

This new upper bound depends on the assumption that the neutron produced by the proton-electron interaction might disobey the exclusion principle, although the parent proton obeyed it. However, the force in this interaction is the weak nuclear force, which is well understood, and appears to give a situation similar to that for electrons in an atom: could an obedient proton give birth to a disobedient neutron? If not, the upper limits remains, for the present, at the level set by Ramberg and Snow.

The violation of the exclusion principle has even been put forward as a solution to the solar-neutrino problem⁶ — the discrepancy between the flux of neutrinos expected and observed from solar fusion reactions (see Michael Cherry's recent News and Views discussion⁷). If protons could violate the principle, two of them could get close enough to be bound together by the strong nuclear force, just as a proton and a neutron can be bound to form a deuterium nucleus. This would increase the rate at which hydrogen burns in the Sun to a level which becomes incompatible with observed properties of the Sun if the probability of violation is more than 1 in 10^{15} . However, Plaga⁶ suggested that a level of violation somewhat lower than this could explain the solar-neutrino problem⁶, as the new non-paulian burning of hydrogen would produce no neutrinos. The amount of violation required is 1 in 10^{18} , which is ruled out by the limits of Kekez *et al.* (but not necessarily by those of Ramberg and Snow, as electrons and nucleons might have different rates of violation). If the objection raised in the previous paragraph is correct, however, Plaga's suggested solution of the solar neutrino problem remains a possibility. □

Tony Sudbery is in the Department of Mathematics, University of York, Heslington, York YO1 5DD, UK.

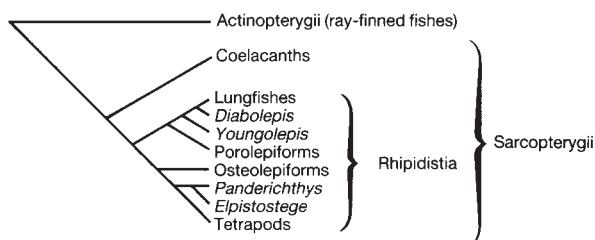
1. Kekez, D., Ljubičić, A. & Logan, B.A. *Nature* **348**, 224 (1990).
2. Ignatiev, A. Yu. & Kuzmin, V.A. *Sov. J. nucl. Phys.* **46**, 444–446 (1987).
3. Greenberg, O.W. & Mohapatra, R.N. *Phys. Rev. Lett.* **59**, 2507–2510 (1987).
4. Govorkov, A.B. *Phys. Lett.* **A137**, 7–10 (1989).
5. Ramberg, E. & Snow, E.A. *Phys. Lett.* **B238**, 438–441 (1990).
6. Plaga, R. *Z. Phys.* **A333**, 397–403 (1989).
7. Cherry, M. *Nature* **347**, 708–709 (1990).
8. Novikov, V.M. *et al. Phys. Lett.* **B240**, 227–231 (1990).

Fossil fishes and fashion

Henry Gee

FASHION-conscious palaeontologists take note: rhipidistians are *à la mode*. The group, which was previously taken to be composed entirely of extinct, lobe-finned fishes, has been resurrected as a clade containing both lungfishes and tetrapods, including ourselves. General agreement over the validity of this arrangement has been brewing for some time, a result of the convergence of views on vertebrate inter-relationships, and the idea received further support at a recent meeting*.

The lobe-finned fishes, or Sarcopterygii,



Simplified cladogram illustrating the consensus view on sarcopterygian interrelationships.

traditionally include lungfishes, coelacanths and rhipidistians (with two main subgroups, osteolepiforms and porolepiforms) as well as tetrapods. As the only wholly extinct sarcopterygian group, rhipidistians were often cast as ancestors of coelacanths and tetrapods. Lungfishes, on the other hand, are so specialized that it was widely assumed that they must have branched off from the other sarcopterygians at a very early stage. Cladistic scrutiny¹ dealt a body blow to both these assumptions. The Rhipidistia were shown to have been characterized (if at all) almost solely on their supposed ancestral relationships with other forms, and thus to lack proper defining characters in their own right. The constituent groups — osteolepiforms and porolepiforms — were better defined, but there seemed little reason to unite them as a larger group. As for the lungfishes, their unique specializations defined them as a natural group, but nothing could be said about their relationships.

In 1981, when Rosen *et al.*¹ united lungfishes with tetrapods as immediate sister groups and consigned the Rhipidistia to symplesiomorphic oblivion, so little was known about many of the fossil lobe-fins that it was hard to see how fossil and Recent groups might be related. But much has changed over the past decade, so much so that the Rhipidistia has every chance of rehabilitation.

First, new fossils have turned up that tie lungfishes and tetrapods to specific yet separate extinct rhipidistian groups. From

China come *Diabolepis* and similar forms^{2,3} which are clearly related to the lungfishes but also show strong similarities with porolepiform rhipidistians. *Panderichthys* and *Elpistostege* (from Latvia and Canada respectively) combine features of osteolepiform rhipidistians with those of tetrapods^{4,5}. These animals, with their dorsoventrally flattened shape, lack of dorsal fins and possession of true frontal bones, are so tetrapod-like that *Elpistostege* was once thought to be a tetrapod when the skull roof was all there was to betray its existence⁶. The true importance of fossils such as *Diabolepis* and *Panderichthys* is that they show subtle derived characters which link the earliest lung-fishes and tetrapods with their immediate relatives, but which have become quite unrecognizable in the modern, specialized forms.

Second, cladistic analyses of the entire sarcopterygian assemblage have shown that some of the old-fashioned rhipidistian characters (such as teeth with folded dentine) are probably derived features after all, not shared by the likes of coelacanths. Far from being descendants of the rhipidistians, coelacanths now appear to have been the first group to have split from the sarcopterygian stem.

Several convincing sarcopterygian phylogenies have now been produced, all of which look very similar thanks to the soundness of the new data^{7,8}. The latest was outlined by Richard Cloutier (Natural History Museum, London) at the meeting (see figure), but Per Erik Ahlberg (University of Oxford) has reached much the same conclusions independently. All of the phylogenies share the same basic features: among the extant groups, lungfishes and tetrapods are more closely related to each other than either are related to coelacanths, but the immediate relatives of both lungfishes and tetrapods are extinct rhipidistians. In cladistic terms, the Rhipidistia can be redefined as a natural group, comprising all the descendants of the common ancestor of lungfishes and tetrapods.

Nevertheless, the case is not yet closed.

1. Rosen, D., Forey, P.L., Patterson, C. & Gardiner, B.G. *Bull. Am. Mus. nat. Hist.* **167**, 163–275 (1981).
2. Chang, M.M. & Yu, X. *Proc. Linn. Soc. New South Wales* **107**, 171–184 (1984).
3. Chang, M.M. thesis, Univ. Stockholm (1982).
4. Vorobyeva, E.I. *Zool. J. Linn. Soc.* **70**, 191–201 (1980).
5. Schultze, H.P. & Arsenault, M. *Palaeontology* **28**, 293–301 (1985).
6. Westoll, T.S. *Nature* **141**, 127–128 (1938).
7. Maisey, J.G. *Cladistics* **2**, 201–256 (1986).
8. Janvier, P. *Océanis* **12**, 123–138 (1986).
9. Yu, X. *J. Vertebrate Palaeontol.* **10** (suppl. to no. 3), 51 (1990).

* 38th Symposium of Vertebrate Palaeontology and Comparative Anatomy, Milton Keynes, UK, 17–21 September 1990.

Jaw fragments from the Upper Devonian of Scotland, under study by Ahlberg and presented at the meeting, may be those of a tetrapod. If so, they would antedate the earliest-known tetrapod body fossils (*Ichthyostega* and *Acanthostega* from Greenland, and *Tulerpeton* from the Soviet Union) by about ten million years. But some of the jaw's features resemble

those of panderichthyids. The truth may lie somewhere in between. And three new genera of porolepiform fishes from the Lower Devonian of China currently under study⁹ promise to add crucial information to our picture of sarcopterygian evolution. □

Henry Gee is an assistant editor of *Nature*.

ANTIGEN PRESENTATION

Naturally processed peptides

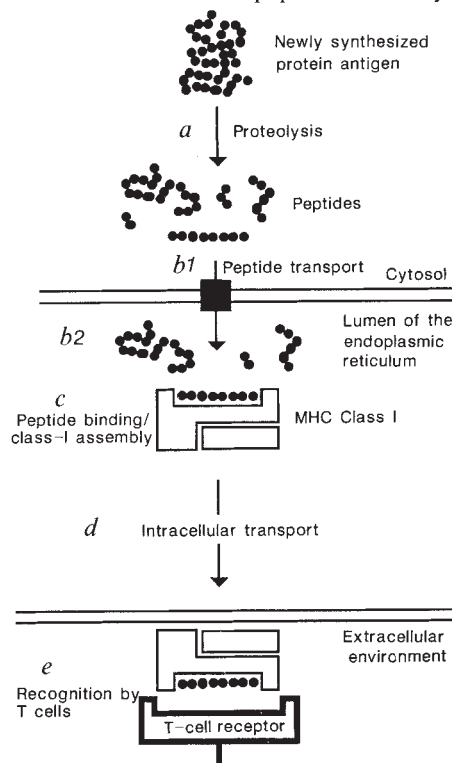
Tim Elliott, Alain Townsend and Vincenzo Cerundolo

THE form in which protein antigens are recognized by class I-restricted cytotoxic T lymphocytes (CTL) has, until recently, been ill defined. There is good evidence that CTL recognize proteolytic fragments of intracellular protein antigens. Most of it has depended on mimicking the naturally processed antigen with synthetic peptides *in vitro*. Now, the ingenious application of high-performance liquid chromatography by two groups has resulted in the identification of the natural fragments of both viral and host antigens presented to CTL¹⁻⁵. Three of the reports³⁻⁵ appear in this issue.

Cytotoxic T lymphocytes recognize peptides derived from intracellular antigens presented at the cell surface in association with class I glycoproteins encoded within the major histocompatibility complex. Various lines of evidence suggest the following events in the presentation of antigens with class I molecules (see figure; for reviews, see refs 6-8). Antigenic proteins are degraded in the cytosol (*a* in the figure), and peptides derived from them are transported into the endoplasmic reticulum (*b1*); it is also possible that proteolysis of proteins in the endoplasmic reticulum might provide additional peptides for binding to class I (*b2*), but as yet there is no evidence for this. In the endoplasmic reticulum, peptides bind to, and take part in, the assembly of class I molecules⁹ (*c*). Unbound peptides may have an extremely short half-life. Finally, the class I/peptide complex is transported to the cell surface (*d*), where it can be recognized by CTL (*e*). Synthetic peptides composed of between 8 and 25 residues can mimic the epitopes recognized by CTL *in vitro* when exposed to the cell surface. But, until now, it was not known whether cells actually produced peptides of this length from intracellular protein antigens.

Peptides have been isolated from whole cells by acid extraction and fractionation of cell extracts either directly¹, or after digestion with a specific endopeptidase². The first experiments showed that minor histocompatibility antigens can, like intracellular viral proteins¹⁰, be identified as peptides; previously they could be defined only by genetic analysis and tissue

graft rejection. An intriguing aspect of the results was that each of the active peptides defined as an intracellular antigen eluted as a single peak (or in one case with a small additional peak) from undigested cell extracts¹. Each of the peptides naturally



Events leading to the presentation of intracellular antigens to major histocompatibility complex class-I restricted T cells. See paragraph 2 of the text for explanation.

produced by the cell was therefore surprisingly homogenous in length.

On page 252, Röttschke *et al.*³ now extend these results to the presentation of influenza virus proteins to class I-restricted CTL. Peptides were isolated from infected cells. The extracts were not treated with a protease, thus allowing identification of only those peptides that arose through the action of the cells' own proteolytic enzymes. Again, the naturally processed peptides eluted as single sharp peaks. This was in stark contrast to the products obtained from a standard peptide synthesis *in vitro*, when minor contaminants of various lengths gave rise to a series

RÉSUMÉ

Timeless gravity

THE strength of gravity has changed by less than one part in a million million each year since the beginning of the Universe, according to new calculations by F. S. Accetta, L. M. Krauss and P. Romanelli (*Phys. Lett. B* **248**, 146-150; 1990). The key to the calculation is, curiously, the nucleosynthesis that occurred in the first four minutes after the Big Bang. Estimates of the cosmic abundances of light isotopes, together with new measurements at CERN of the number of types of neutrinos and improved measurements of the radioactive half-life of the neutron, are used to constrain calculations of the rate at which the Universe expanded in these first few moments — in turn determining the strength of gravity at the time. Although the new result still allows for a 40 per cent variation in Newton's gravitational 'constant' ('big G') since the Big Bang, this hardly reflects the scale of change envisaged by P. A. M. Dirac 50 years ago (*Nature* **139**, 323; 1937).

Staying cool

ELEPHANTS have a problem: disposing of the heat from their imposing bodies. Using infrared thermography, T. M. Williams has mapped the skin temperature for an adult African elephant and an immature Indian elephant, thus determining the main routes by which they lose heat (*J. Zool. Lond.* **222**, 235-245; 1990). At an ambient temperature of 12.6 °C, 86 per cent of total heat loss is accounted for by free convection and radiation. Surprisingly, only 8 per cent of the total heat transfer occurred across the African elephant's ears, traditionally taken to be the main route for heat dissipation. However the thermal conductance of both species of elephant is 3-5 times higher than that predicted from allometry, which Williams attributes to the absence of fur, an adaptation which clearly helps them to stay cool.

Bye cycles

DON'T do it, is the advice of A. K. Baksi to those who would seek evidence of cyclical events in Earth history. Writing in *Geology* (**18**, 983-986; 1990), Baksi discusses various hypotheses proposing that events have occurred with periodicities of around 30 million years (flood-basalt volcanic episodes, cometary impacts and faunal extinctions) and finds all to be flawed. The radiometric dates used, he says, are quite simply too unreliable and were questionably selected. "Further search for periodicity in global events," he adds, "should be postponed until an adequate data base of high-quality radiometric ages becomes available." But recent advances in analytical techniques should allow quick progress on this front.

of peaks, many of which could be recognized by CTL. The result emphasized the homogeneous nature of the peptides isolated from both infected and uninfected cells, and allowed Rötzschke *et al.* to identify the synthesized sequences that co-eluted with the peptide isolated from infected cells. In both cases the naturally processed peptide was shorter and more potent than the original synthesized sequences used to define the epitopes *in vitro*.

In the original experiments that defined the D^b-restricted epitope from influenza nucleoprotein (NP), the sequence of residues 366–379 was the most active of a set of NP peptides that were progressively shortened at their N termini¹¹. The naturally presented sequence is the NP nonamer 366–374. Similarly, the NP peptide 147–158 defined the K^d-restricted epitope¹² of NP, which contains the naturally processed peptide identified by Rötzschke *et al.* as the NP nonamer 147–155.

So, of all the possible peptides of varying length that could be presented by class I molecules and recognized by CTL, the naturally presented peptide epitope is a single species, composed only of eight or nine amino-acid residues. In contrast, peptides extracted from purified class II molecules seem to be heterogeneous in length¹³. The difference may result either from a fundamental difference in the binding sites of class I and II molecules, or

from the way in which protein antigens are degraded and transported before presentation by class I and II.

In another report (page 248), Falk *et al.*⁴ show that the peptide defining any particular antigen can be isolated only from cells that express the class I molecule known to function as a restriction element for that antigen. The class I molecules in a cell therefore define the set of peptides that can be isolated by acid extraction. Although this observation has yet to be controlled by comparing cell lines that differ only by the expression of a transfected class I gene, the evidence from congenic and class I mutant strains of mice implies that the class I molecules themselves are responsible for this selective effect.

How then can the class I molecules define the pool of peptides within a cell? Theoretically, class I molecules could be involved either in peptide production, or in protection of peptides from degradation to amino acids. Falk *et al.*⁴ offer a variety of models along these lines (Fig. 3 of their paper). It is difficult to see how class I molecules could be involved in the generation of peptides in the cytosol, because class I is thought to be cotranslationally transported into the endoplasmic reticulum. Given the evidence that class I molecules can bind peptides of a variety of lengths *in vitro*, perhaps, as suggested by Falk *et al.* they can also bind a variety of

peptides that, in the endoplasmic reticulum *in vivo*, differ in length. These peptides could be generated either by transport from the cytosol or by degradation of larger proteins in the endoplasmic reticulum. Once bound, peptides may be rapidly trimmed by proteases until they reach a length that fits snugly in the binding site and are then protected from further degradation. Unbound peptides may be degraded rapidly to single amino acids.

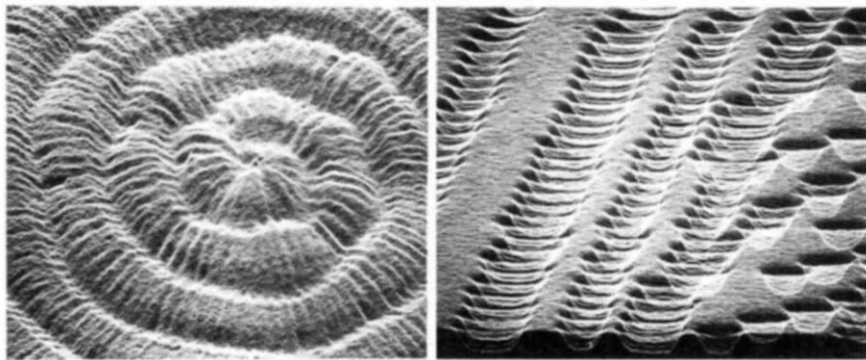
The dimensions of the antigen-binding sites of HLA A2 and Aw68 could, in theory, accommodate peptides consisting of 8–20 amino acids, depending on how compressed the peptide might be by folding^{14,15}. In the structures of both of these class I molecules, peptide-like material could be seen in the binding site, but could not be defined, suggesting that naturally processed peptides could be isolated from purified class I molecules. Indeed, radioiodinated peptides have been extracted from purified HLA A2 molecules, and fell within a narrow range of molecular weights¹⁶.

Van Bleek and Nathenson⁵ (see page 213) have now applied the same fractionation technology to purified K^b class I molecules from cells infected with vesicular stomatitis virus (VSV), with the aim of isolating viral peptides bound to K^b. They have succeeded in identifying a peptide derived from the VSV nucleocapsid (N) protein, which copurified with K^b but not D^b molecules. This peptide eluted as a single peak, and was an octamer representing residues 52–59. The result offers a clear interpretation of the experiments done on total cell extracts and strongly suggests that binding of peptide by the class I molecule is the basis for the selection of peptides that can be isolated by acid extraction.

We have suggested that the antigen presentation/class I assembly mutants RMA-S and .174/T2 have a defect in signal-independent peptide transport from the cytosol to the endoplasmic reticulum^{9,17}. Peptides originating in the cytosol (such as those derived from influenza nucleoprotein) may be degraded completely if not transported immediately and protected by a class I molecule. Then, despite the presence of class I molecules in these cells, peptides originating from the cytosol may not be detectable by acid extraction. This result would also show that the presence of class I molecules within a cell is not by itself sufficient to influence the extractable peptide pool.

The uniform length of the peptides isolated from class I molecules may simply reflect the size of the class I binding site, and efficient N- and C-terminal trimming of the bound peptide, as suggested by Falk *et al.*⁴. But this presupposes that all peptides bound in the site are in extended conformation and are not folded (and

Photon tunnelling microscopy



SCANNING tunnelling microscopes are now almost routine devices that can trace the topography of a surface through the tiny chance that an electron can leap the insulating gap between two closely separated conducting surfaces, one a nanometre-sized microscope stylus, the other the sample. The analogous optical device uses the related effect, that photons can tunnel across a similar narrow gap. As John M. Guerra points out in a special issue of *Applied Optics* on three-dimensional microscopy (29, 3741–3752; 1990), photon tunnelling is just a new description of an effect known since Newton's time — frustrated total internal reflection. Light reflected at a glancing angle from a glass/air interface (as in a prism) penetrates just past the glass surface to form an 'evanescent wave'. A material placed within this evanescent wave, less than a wavelength from the glass surface, disturbs it and frustrates the total internal reflection. The micrographs above were prepared using this more conventional form of microscopy, but combined with modern fast electronics and image processing developed for the scanning microscopes to generate three-dimensional pictures of sample surfaces — diamond-turned acrylic (left) and compact-disk pits (right). Guerra notes that the method combines the extreme vertical resolution (but not the horizontal resolution) of the scanning microscopes with the convenience of the conventional near-field microscopes which image whole fields in a single go.

R.P.

therefore compressed) in any way. An alternative explanation is that some factor other than the dimensions of the class I binding site controls the length of peptide isolated by acid extraction. One possibility is that the mechanism which transports peptides from the cytosol to the endoplasmic reticulum, and which provides peptides for binding to class I, is selective for sequences of 8 or 9 amino acids.

Peptide transport is independent of hydrophobic signal sequences, and is therefore distinct from the classical vectorial transport of secreted or membrane proteins (for reviews, see refs 6 and 7). Although there is no clue as to the molecular basis of peptide transport, there is a precedent for transport of peptides across bacterial membranes in the form of oligopeptide permease, which may be selective for sequences of 3–5 amino acids. This molecule is a member of the 'ABC' (ATP-binding cassette) transporters, which have several counterparts in eukaryotes, including the product of the yeast *STE6* gene that transports a defined peptide sequence (for reviews, see refs 18 and 19). So it is conceivable⁶ that there is a peptide transporter which selects peptides of particular lengths for translocation across the membrane of the endoplasmic reticulum, and which is defective in the mutants RMA-S and .174/T2. Finally, it may be that controlled proteolysis in the cytosol gives rise to peptides with an average length of 8 or 9 amino acids. By isolating a large number of naturally presented peptides and looking for common features of their N and C termini, it may be possible to identify elements associated with specific proteolysis or transport. □

Tim Elliott, Alain Townsend and Vincenzo Cerundolo are in the Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK.

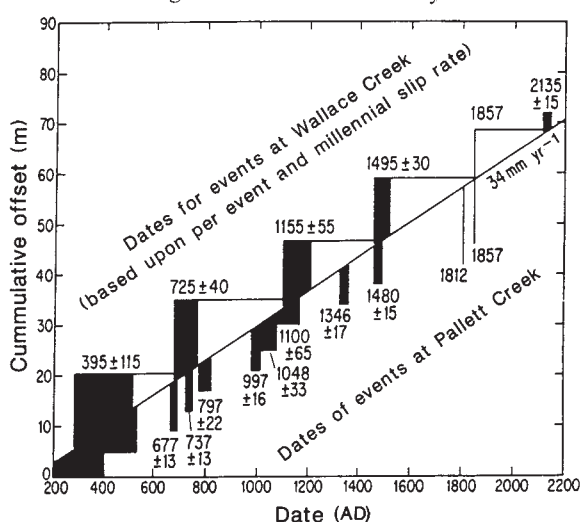
1. Röttschke, O., Falk, K., Wallny, H. J., Faath, S. & Rammensee, H.-G. *Science* **249**, 283–287 (1990).
2. Wallny, H. J. & Rammensee, H.-G. *Nature* **343**, 275–278 (1990).
3. Röttschke, O. *et al.* *Nature* **348**, 252–254 (1990).
4. Falk, K., Röttschke, O. & Rammensee, H.-G. *Nature* **348**, 248–251 (1990).
5. Van Bleek, G. M. & Nathanson, S. G. *Nature* **348**, 213–216 (1990).
6. Townsend, S. & Bodmer, H. A. *Rev. Immun.* **7**, 601–624 (1989).
7. Yewdell, J. W. & Bennink, J. R. *Cell* **62**, 203–206 (1990).
8. Braciale, T. J. *et al.* *Immun. Rev.* **98**, 95–114 (1987).
9. Townsend, A. *et al.* *Nature* **340**, 443–448 (1989).
10. Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457–467 (1985).
11. Townsend, A. R. M. *et al.* *Cell* **44**, 959–968 (1986).
12. Taylor, P. M., Davey, J., Howland, K., Rothbard, J. & Askonas, B. A. *Immunogenetics* **26**, 267–272 (1987).
13. Buus, S., Sette, A., Colon, S. & Grey, H. *Science* **242**, 1045–1047 (1988).
14. Bjorkman, P. J. *et al.* *Nature* **329**, 512–516 (1987).
15. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692–696 (1989).
16. Elliott, T. J. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5213–5217 (1990).
17. Cerundolo, V. *et al.* *Nature* **345**, 449–456 (1990).
18. Higgins, C. F. *et al.* *J. Bioenerg. Biomemb.* **22**, 571–593 (1990).
19. Higgins, C. *Nature* **341**, 103 (1989).

Earthquakes as chaos

Christopher H. Scholz

EARTHQUAKE hazard analysis has for some time been based on the supposition that faults are broken into discrete segments that independently fail in 'characteristic' earthquakes. Long-range predictions can then be made based on the identification of these segments and estimating their recurrence times. It is becoming increasingly obvious that this description is far too simple. One problem is that the segments are not entirely

Although impossible to prove with radiocarbon dates, it is reasonable to suppose that the Wallace Creek events are coupled events like that of 1857, in which both sites were ruptured. If the coupled events were the last in each Pallett Creek sequence, then in detail the behaviour is astonishingly like Huang and Turcotte's model, in which Wallace Creek corresponds to the stronger block and Pallett Creek the weaker. After each coupled event, there is



The history of slip in earthquakes at two sites about 200 km apart on the San Andreas fault (from ref. 2).

independent: two or more adjacent segments sometimes rupture together to produce a much larger earthquake than would otherwise have been expected. On page 234 of this issue¹, Huang and Turcotte provide a crisp explanation of this phenomenon. They have constructed a model in which two sliders with different friction are coupled together by one spring and driven at the same rate by another. They find that this system evolves into chaos, in which the two sliders exhibit independent stick-slip oscillations for some periods but sometimes couple together to slip in a larger event.

An example of this behaviour can be seen in the history of slip in earthquakes at two sites, Wallace Creek and Pallett Creek, some 200 km apart on the San Andreas fault (see figure) which has been deduced for the past 2,000 years using techniques akin to those used in archaeology². The last great earthquake to have ruptured this section of the fault in 1857 ruptured through both of these sites, slipping 9 m at Wallace Creek and 2 m at Pallett Creek, values of slip that appear characteristic of these sites for the prehistoric earthquakes as well. Earthquakes at Pallett Creek occur in clusters of two or three events, each of which corresponds in time to one event at Wallace Creek.

a long quiescence at Pallett Creek followed by a series of more frequent events, culminating in another coupled event, just as this model predicts (see their Fig. 3, on page 235 of this issue).

Other interesting cases have yet to be evaluated. The Loma Prieta earthquake has refocused attention on seismic hazard in the San Francisco Bay Area. One intriguing question is the possibility of coupling between the Hayward and San Andreas faults on either side of San Francisco Bay. In 1836 there was a magnitude-7 event on the northern half of the Hayward fault, followed two years later by a similar

San Andreas earthquake on the San Francisco Peninsula. In 1865 a magnitude-6.5 event in the vicinity of Loma Prieta was followed three years later by a magnitude-7 event on the southern Hayward fault. An analysis of this problem with Huang and Turcotte's model should be revealing.

The general problem appears to be considerably deeper than these examples indicate. Thatcher³ has shown that the rupture segments are not stationary in time: their shape and spatial distribution on the fault change from earthquake cycle to cycle (see Max Wyss's recent News and Views article⁴). Furthermore, the recurrence of rupture on any fault sector is now known not to be periodic but lognormally distributed⁵. These factors indicate behaviour that is far too complex to be described by the low-dimensional chaos of the type studied by Huang and Turcotte, but progress is being made on this front as well.

The model of Bak *et al.*⁶ of a spatially extended dissipative system which evolves to a self-organized critical state is a very attractive general model of earthquakes, as they have pointed out⁷. It predicts that earthquakes obey a fractal size distribution — a distribution known historically in seismology as the Gutenberg–Richter

relation. But this is not the whole story. For good scientific reasons, Bak *et al.* were careful to evaluate the behaviour of their model far from its boundaries, to avoid any spurious effects arising therefrom. It now seems that this missed an essential feature of earthquake behaviour.

Faults slip unstably in earthquakes only in a narrow depth range between the surface and a depth, commonly about 15 km, at which the rocks become ductile. Thus the problem contains a characteristic dimension not included by Bak *et al.* in their idealization. It has been known for some time that earthquakes of size smaller than this dimension are fundamentally different from those larger, whose rupture is then constrained to be channelled horizontally. 'Small' earthquakes, so defined, obey different scaling laws from 'large' earthquakes^{8,9}. Furthermore, although both small and large earthquakes have fractal size distributions, they belong to disjoint sets so that for a given section of fault many more large earthquakes will occur than predicted by extrapolating from small events¹⁰.

Carlson and Langer¹¹ studied a model with spring slider elements as in Huang and Turcotte's model but which was spatially extended like the Bak *et al.* model. They found that this model also exhibits self-organized criticality and the Gutenberg-Richter law, but owing to what was first thought to be a weakness of their model — its finite size — they observed the 'artefact' that many more large events were produced than expected from the distribution of small events. Following up this serendipitous clue, my colleagues and I studied a model in the form of a narrow strip that deliberately emphasizes the effect of the lateral boundaries^{12,13}. This model produces two populations, of small and large events. The large events (running from boundary to boundary) are much more frequent than predicted from the distribution of small events. Concentrating on the large events, we found that the model also exhibits two other primary features of natural seismicity: the large events greatly dominate the moment

release rate of the fault, and the recurrence time of large events is lognormally distributed.

In the rather short time since this approach to earthquakes was first discussed at a seminal meeting in 1989 (see my report in News and Views¹⁴), progress has been very rapid. Whereas the earthquake process is now being viewed as far more complex than before, there is clearly

a breakthrough in the making in understanding the underlying physics. It is too early to say what effect this will have on practical matters such as earthquake prediction. □

Christopher H. Scholz is in the Lamont-Doherty Geological Observatory and the Department of Geological Sciences, Columbia University, Palisades, New York 10964, USA.

PROTEIN CONFORMATION

Hinge-bending and folding

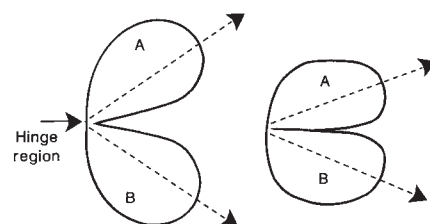
Christopher M. Dobson

THE idea that conformational transitions in proteins can involve the relative movement of essentially rigid structural elements is an attractive one, not least because of its simplicity and ready visual impact. Particularly well-known examples are the transitions of allosteric proteins, such as haemoglobin¹ or phosphorylase², where individual subunits move relative to each other in a cooperative manner. Within single subunit proteins, a related and intriguing concept has been that of 'hinge-bending', whereby the relative flexibility of short regions of the polypeptide chain allows significant movement of structural regions, or domains, within the cooperative folding unit. X-ray diffraction studies by Faber and Matthews, reported on page 263 of this issue³, provide a striking illustration of this idea, as well as demonstrating the effects of interactions within crystals on the selection of particular conformational states of proteins.

Faber and Matthews describe experiments with a mutant form of lysozyme from the T4 bacteriophage. T4 lysozyme is an enzyme of 164 amino-acid residues, whose structure consists of two lobes (or domains) of approximately equal size. Although having no apparent sequence similarity, its structure is related to that of the avian and mammalian c-type lysozymes of which that from hen egg white is the best known. It has long been advocated that hinge-bending occurs in egg-white lysozyme⁴, but experimental verification has proved elusive⁵.

The mutant T4 lysozyme (Met 6 → Ile) crystallizes in a trigonal space group which is isomorphous with the wild-type protein; the structure in this form is very like that of the wild type. Faber and Matthews found, however, that the mutant lysozyme also crystallizes in an orthorhombic space group, containing four molecules in the asymmetric unit. Solving this structure by isomorphous replacement proved to be more complicated than expected, because the relative orientation of the domains of the lysozyme molecules in the orthorhombic crystals differs significantly from that found in the structures of the wild-type and mutant protein in the trigonal

crystals. Moreover, it differs substantially between the four molecules within the asymmetric unit. The domains themselves are closely similar in each case, but the hinge-bending angle varies by up to 32°. Despite the large 'global' changes in struc-



A schematic example of one concept of hinge-bending. A protein structure may consist of distinct structural 'domains' (A and B in the diagram above), linked together by a segment, or segments, of the polypeptide chain, denoted the hinge region, in such a way that the relative orientation of the domains can vary without significant structural changes in the domains themselves. The domains may be composed of continuous regions of the polypeptide chain, as in T4 lysozyme where the N- and C-terminal regions each fold to distinct domains. They may also arise from the association of discontinuous regions of the sequence; examples are the c-type lysozymes which, although topologically related to T4 lysozyme, have one domain consisting of the N- and C-terminal region together, with the region in the centre of the sequence forming the other domain. The relative movement of the domains can cause large global conformational changes which may permit access of substrates and generate an appropriate environment for catalysis.

ture giving rise to a relative movement of certain residues of close to 10 Å, the conformational changes at the hinge-bending site turn out to be small.

The conformational flexibility offered by hinge-bending is of particular note in that it may be important in allowing access of substrates to the active site, which is formed at the interface of the two domains. Indeed, there are several well-documented cases of such structural transitions resulting from the binding of ligands to proteins (hexokinase and phosphoglycerate kinase being particularly clear examples⁶). What distinguishes the Faber and Matthews result is that it seems

- Huang, J. & Turcotte, D.L. *Nature* **348**, 234–236 (1990).
- Sieh, K., Stuiver, M. & Brillinger, D. *J. geophys. Res.* **94**, 603–623 (1989).
- Thatcher, W. *J. geophys. Res.* **95**, 2609–2623 (1990).
- Wyss, M. *Nature* **345**, 290 (1990).
- Nishenko, S. & Buland, R. *Bull. seismol. Soc. Am.* **77**, 1382–1399 (1987).
- Bak, P., Tang, C. & Wiesenfeld, K. *Phys. Rev. Lett.* **59**, 381–384 (1987).
- Bak, P. & Tang, C. *J. geophys. Res.* **94**, 15635 (1989).
- Scholz, C. *Bull. seismol. Soc. Am.* **72**, 1–14 (1982).
- Shimazaki, K. in *Earthquake Source Mechanics* (eds Das, S., Boatwright, J. & Scholz, C.) (Am. Geophys. Un., Washington DC, 1986).
- Davison, F. & Scholz, C. *Bull. seismol. Soc. Am.* **75**, 1349–1362 (1985).
- Carlson, F. & Langer, J.S. *Phys. Rev. Lett.* **62**, 2632–2635 (1989).
- Brown, S., Rundle, J. & Scholz, C. *Eos* **71**, 467 (1990).
- Brown, S., Rundle, J. & Scholz, C. *Geophys. Res. Lett.* (in the press).
- Scholz, C.H. *Nature* **338**, 459–460 (1989).

to be the first example of a crystal structure in which such large conformational variability is observed within a highly cooperative folding unit which is not in response to ligand binding. Much smaller variability has been reported, such as that involving individual helices of insulin in different crystallographic environments⁷, and such 'helix-shear' processes involving rigid-body-like motions have been proposed as a means of transmitting conformational change across protein structures⁸. But although there have been observations of large domain variability within single subunit proteins, notably the immunoglobulins⁹ and a viral coat protein¹⁰, Faber and Matthews argue that such domains, unlike those of T4 lysozyme, are essentially independent folding units.

One question that arises is whether the differences in hinge-bending observed in the mutant T4 lysozyme reflect the distribution of conformations found in solution, or whether they are generated by interactions within the crystal. As the present structural analysis, along with previous theoretical calculations¹, suggest that a hinge-bending mode is likely to be a low-energy deformation, the former could well be the case. Certainly, Faber and Matthews favour this view and cite the temperature factors of the wild-type structure¹¹ as supporting evidence. It will be interesting to see whether other techniques can provide direct information about the solution state. In the case of some proteins which have domains that are independent folding units, conformational variability has, for example, been evident from dynamic behaviour detected in NMR experiments¹².

Interest in hinge-bending has recently been heightened by findings which suggest that the phenomenon may not in fact be confined to well-defined domains of proteins. In studies of triose phosphate isomerase by Joseph, Petsko and Karplus¹³, binding of substrates has been found to induce a conformational transition involving an 11-residue loop which closes over the active site; in the absence

of substrate, the loop projects into the solvent. From comparison of the structures of the open and closed forms of the protein, and molecular dynamics calculations, it would seem that the loop moves in a rigid-body type of displacement analogous to the hinge-bending motion used to describe domain motions.

The possible significance of the concerted motions of such small regions of proteins has also been raised in a different context. In a study of foot and mouth disease virus, Parry *et al.* observed at least two conformations, or families of conformations, of a 17-residue loop of coat protein¹⁴. The difference in this loop region between variant viruses with differing antigenic properties implies that the conformational change can destroy the integrity of the epitope and hence offer a novel mechanism for the virus to circumvent antibody binding. Whether or not this can be related to hinge-bending remains to be seen.

Faber and Matthews's results represent a clear case of the ability of domains in a cooperative folding unit to maintain their integrity whilst undergoing considerable

changes in their relative orientation. Could this be significant not only for the behaviour of the folded protein but for the folding process itself? It has, for example, been suggested that regions within cooperative protein structures can form in the folding process independently in the absence of other regions of the structure, and that this has a large part in determining folding pathways. Experimental justification for such a view is increasingly evident, for example from studies of carefully designed peptides and of the increasingly well-characterized 'molten globule' states of proteins, as well as from structural studies of intermediates generated during the refolding of proteins. These issues have been discussed in a recent News and Views¹⁵. Results of the type described by Faber and Matthews, and the light they shed on domain interactions, may therefore turn out to be of even wider significance than at first meets the eye. □

Christopher M. Dobson is in the Oxford Centre of Molecular Sciences and Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, UK.

OCEAN DYNAMICS

Not so quiet on the ocean front

Robert A. Weller

OCEAN fronts are regions of strong horizontal gradients in water properties. In the open ocean they are, in effect, rather narrow boundary regions between bodies of water characterized by different temperature or salinity, or some other differing property. Since the 1960s we have known¹ that jet-like, along-front surface currents with speeds of up to 1 m s⁻¹ are found in association with the strong cross-frontal gradients. Pollard² and others have more recently described considerable structure and variability in the upper 300 m in the vicinity of fronts. This variability in the density and flow fields was found to be dominated by mesoscale (100 km) and smaller-scale eddies. On page 227 of this issue³, Pollard and Regier show that small-scale regions, tens of kilometres across, result from the eddies and the distorted front, in which vertical velocities at the base of the nearly isothermal mixed layer may be as large as tens of metres a day. This is interesting not only to those studying the dynamics of the upper ocean, but also to those trying to explain the mesoscale and smaller-scale patchiness found in phytoplankton populations and hoping to elucidate the vertical exchange of gases and other constituents of seawater from near the surface to the interior of the ocean.

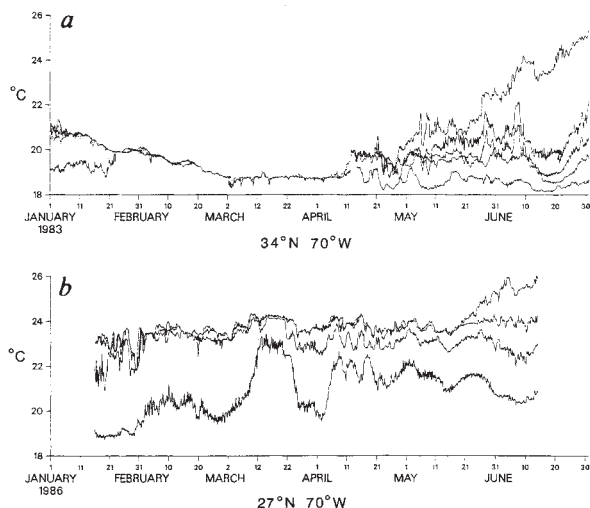
In the western North Atlantic, where Pollard and Regier made their observations, the Westerly Winds force a south-

ward transport of water from the Sargasso Sea towards the region, roughly between 25° N and 29° N, where ocean fronts are common. At the same time, the Trade Winds force a northwards flow of water from the subtropics towards the frontal region. In these source regions, for example, in the Sargasso Sea, the vertical structure of the upper ocean (*a* in the figure) is largely determined by local atmospheric forcing. In the winter, heat is lost to the atmosphere and surface water cools, becomes more dense and sinks. As the cooling persists, a relatively deep (hundreds of metres) layer of nearly constant temperature forms.

In spring and summer, the net heat flux changes sign and the ocean gains heat. At the same time, the winds moderate; and the combination of heat gain and reduced wind-driven mixing allow the upper ocean to warm and a shallower mixed layer to form, perhaps tens of metres deep. This shallow layer is isolated from the interior by its buoyancy. Vertical exchange between the mixed layer and the fluid below will not occur again until storms and the winds of the autumn cause entrainment of fluid from below and the deepening and cooling of the mixed layer.

One result of the isolation of surface water from the interior is that the nutrients within the surface layer can be consumed and phytoplankton populations become limited by nutrient concentra-

1. Perutz, M.F. *Nature* **228**, 727–734 (1970).
2. Barford, D. & Johnson, L.N. *Nature* **340**, 609–616 (1989).
3. Faber, H.R. & Matthews, B.W. *Nature* **348**, 263–266 (1990).
4. McCammon, J.A., Gelin, B.R., Karplus, M. & Wolynes, P.G. *Nature* **262**, 325 (1976).
5. Sternberg, M.J.E., Grace, D.E.P. & Phillips, D.C. *J. molec. Biol.* **130**, 231–253 (1979).
6. Janin, J. & Wodak, S. *Prog. Biophys. mol. Biol.* **42**, 21 (1983).
7. Chothia, C., Lesk, A.M., Dodson, G.G. & Hodgkin, D.C. *Nature* **302**, 500–505 (1983).
8. Levine, B.A. *et al. Ciba Fdn Symp.* **93**, 72–97 (1983).
9. Huber, R. *et al. Nature* **264**, 415–420 (1976).
10. Harrison, S.C., Olgen, A.J., Schutt, C.E., Winkler, F.K. & Bricogne, G. *Nature* **276**, 368–373 (1980).
11. Weaver, L.H. & Matthews, B.W. *J. molec. Biol.* **193**, 189–199 (1987).
12. Oswald, R.E. *et al. Nature* **337**, 579–582 (1989).
13. Joseph, D., Petsko, G.A. & Karplus, M. *Science* **249**, 1425–1428 (1990).
14. Parry, N. *et al. Nature* **347**, 569–572 (1990).
15. Baldwin, R.L. *Nature* **346**, 409–410 (1990).



Time series of temperature at various depths in the upper ocean from two locations at the same longitude in the western North Atlantic. The times series (a) collected at 34° N, 70° W are from within the Sargasso Sea, while those collected at 27° N, 70° W (b) are from the frontal region surveyed by Pollard and Regier. Data in a come from 5, 50, 75, 100, 150 and 200-m depths; note that the upper 200 m is essentially isothermal from late January to mid-April when it warms and stratifies rapidly. In contrast, b, which has data from 10, 40, 80, 120 and 160-m depths, shows an upper ocean where the passage of ocean fronts (such as on about 12 March and again on 22 March) contributed abrupt changes in temperature and where a deep isothermal mixed layer did not form in the winter.

tions alone. On the other hand, if the mixed layer is deep and the vertical circulation active, plankton can be carried below the depth to which light penetrates and their population become limited by light levels. Because the spatial scales of the atmospheric forcing fields (wind, air temperature, relative humidity and radiation) that drive the seasonal cycle in the upper ocean are large, characteristically little less than 1,000 km, there is no reason to expect mesoscale and small-scale spatial variability away from fronts in the mixed-layer depth or in other fields (such as the phytoplankton population) dependent on vertical exchange across the base of the mixed layer. Indeed, the planktonic spring blooms yield evidence of large spatial scales linked to those scales of the atmospheric forcing, a general poleward migration of the near-surface maximum in chlorophyll-*a* concentration coupled to the northward progression of the restratification of the upper ocean in spring⁴.

Yet, in some parts of the ocean, there is considerable variability in chlorophyll content on 10–100 km scales⁵. A link between productivity and mesoscale and small-scale physical processes that carry nutrient-rich water up into the mixed layer have been proposed as an explanation⁶. If such an explanation is to be plausible, the physical processes should be apparent in detailed studies, such as Pollard and Regier's, of ocean regions characterized by energetic mesoscale and smaller scale variability. This is the case. The frontal region the authors surveyed had eddies

embedded in the flow field associated with the front. The front itself was characterized by light (warm) water to the south and heavy (cold) water to the north. Constant density-surfaces sloped upwards from beneath the warm side, pinched together, and reached the ocean's surface with an almost vertical inclination at the front.

Pollard and Regier explain that large, 50-km eddies distorted the frontal flow field, giving rise to smaller eddies, both cyclonic and anticyclonic, which were advected horizontally. As the small eddies moved through changes in stratification and horizontal shears, they conserved the physical quantity termed potential vorticity (the product of the rate of rotation or total vorticity of the water per unit height, times the vertical gradient of its density). This conservation was accomplished by changes in thickness so that, depending on the sign of change in the

fluid's vorticity, the strata were either shrunk or stretched. Shrinking in height would bring water up from below, while stretching to conserve potential vorticity would push near-surface fluid to greater depth. Pollard and Regier calculate that the vertical velocities at the base of the small eddies could be as high as 50 m d⁻¹.

In the area studied by Pollard and Regier, ocean fronts are common. Satellite images showed a number of fronts in the region at any one time, and at one location within the frontal region their occurrence dominated records of near-surface temperature (b in the figure). Further, Pollard and Regier conclude that the eddies, which individually persisted for at least several days, were characteristic features of the fronts. Thus, one conclusion to be drawn from their study is that ocean frontal regions are sites of persistent, active vertical exchange on small horizontal scales (around 10 km) between the near-surface and the upper thermocline regions of the ocean. □

Robert A. Weller is in the Physical Oceanography Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

1. Voorhis, A. D. *J. geophys. Res.* **16**, 3809–3814 (1969).
2. Pollard, R. T. *Nature* **323**, 433–435 (1986).
3. Pollard, R. T. & Regier, L. *Nature* **348**, 227–229 (1990).
4. Strass, V. & Woods, J. D. in *Toward a Theory on Biological-Physical Interactions in the World Ocean* (ed. Rothschild, B. J.) 113–136 (Kluwer, Dordrecht, 1988).
5. Gower, J. F. R., Denman, K. L. & Holyer, R. J. *Nature* **288**, 157–159 (1990).
6. Woods, J. in *Toward a Theory on Biological-Physical Interactions in the World Ocean* (ed. Rothschild, B. J.) 7–38 (Kluwer, Dordrecht, 1988).

Sticking in orbit

SPACE is gradually filling up with junk. The final stage rocket of every satellite is junk immediately; the satellite itself will be junk in due course; rockets that explode and spacecraft that break up spray further enormous numbers of tiny objects into orbit. At orbital velocity, impact with even a tiny piece of this space junk can wreck a satellite. Every new launch increases the danger.

This mass of fragments all in separate orbits, says Daedalus, is rather like the primitive solar system. Why did that ancient distribution of small particles clump together into the big planets we know today, instead of fragmenting further by repeated collisions? Solar system theorists reckon that the primordial planet stuff was *sticky*. Repeated impacts would then have glued it into a few sticky proto-planets big enough to sweep up and capture the remaining dust and gas.

So DREADCO's paint department is formulating a sticky paint for spacecraft. It will simply swallow up any tiny fragment that hits it, and flow over the scar to re-establish a smooth sticky surface. Even if the heat of impact degrades the paint in the immediate vicinity, the resulting charred volume will still be retained by the surrounding paint. If splashed by major impact, the paint is designed to extend like saliva into tenacious strings, which will ultimately contract under surface tension and return to the satellite.

New satellites can simply be painted with the novel coating before launch. Existing satellites and fragments pose a trickier problem. Daedalus recalls his previous scheme for vast space 'soap bubbles', made from a liquid loaded with high-polymer ring molecules of large radius. Such a bubble cannot be punctured. Any small hole is bound to be surrounded by at least one ring-molecule, whose random brownian wriggling soon seals the hole again. A few kilometre-sized sticky space bubbles, says Daedalus, would soon spread their adhesive around. They would simply absorb the very smallest particles, or slow them as they passed through the bubble walls and the gas between, dropping them from orbit. But larger chunks and complete satellites that hit a bubble would be comprehensively coated with adhesive by the impact. Even if they broke the bubble up, glue would be usefully spread round its orbit, ready to stick to further space junk.

The original processes of planet formation will thus be instructively replicated in earth orbit. All the space junk will gradually clump together. Ultimately a few large, messy satellites will result, each one a crazy glued-together agglomeration of space fragments large and small, easily tracked and avoided. The peril will be over!

David Jones

Conclusions from Lake Nyos disaster

SIR—In August 1986, the Lake Nyos disaster brought death and devastation to a densely populated part of the Northwest Province of Cameroon^{1,2}. The death toll, which was officially estimated at more than 1,700, was far in excess of that for any previous similar disaster. In the immediate aftermath many countries not only offered humanitarian assistance but also sent teams of scientists to investigate the cause of the disaster. All participants were invited to attend a conference organized by the Cameroonian Ministry of Higher Education and Scientific Research in Yaoundé in March 1987 to "find out as much as possible about the cause of the Lake Nyos disaster" and "determine the precautions which should be taken to avoid a similar tragedy in the future".

Although the scientific discussions in Yaoundé were interesting, it soon became apparent that there were irreconcilable differences between various groups concerning almost every aspect of the disaster. The final communiqué, which was widely circulated and subsequently published³, accurately reflected the scientific arguments but offered little sound advice to the government of Cameroon other than to suggest that more scientific studies were needed. During the meeting, a group of us set up an international working group to act as an information exchange and coordinate future research⁴.

A seminar on the Lake Nyos gas disaster was convened on 11 September by our group to which all those who had worked on the disaster were invited. During the discussions it became apparent that a consensus had now arisen concerning the cause of the disaster and the urgent need for remedial measures to prevent a recurrence. Our main objective is to offer helpful scientific advice to the political authorities in Cameroon, but we hope that our four main conclusions, which are set out below, will also be of wider interest.

1. The 1986 disaster was caused by a massive release of CO₂ from Lake Nyos.
2. Lake Nyos now contains about 300 million cubic metres of carbon dioxide and therefore remains very dangerous.
3. Further CO₂ is being added at a rate of at least 5 million m³ per year and the danger is therefore increasing.
4. Another gas disaster could occur at any time, so the CO₂ in Lake Nyos should be reduced as a matter of urgency.

The members of the working group considered ways in which carbon dioxide might be extracted from Lake Nyos.

There was general concern that any attempt to remove gas from the lake might itself create an unstable situation and induce a massive gas release. Before any attempt at controlled degassing, therefore, equipment should be installed to monitor the stability of the lake, and appropriate plans should be prepared to deal with any medical emergency which might arise. It is also recommended that any system of gas extraction should first be tested on Lake Monoun, 100 km southeast of Nyos. This lake was the source of a gas release which killed 37 people in 1984 (ref. 5), but as it is only half as deep as Lake Nyos it is potentially far less dangerous.

There was general agreement that pipes should be installed to remove the gas-rich bottom water from the lakes and that the degassed water should be discharged outside the lake basins, so as to avoid disturbing the natural stratification⁶ of the lakes. Concern was expressed over the total lack of information concerning the amount of

gas in the sediments and rocks beneath the lakes. If there are zones of gas-saturated (or near saturated) water beneath either of the lakes, then any reduction in pressure might trigger another lethal gas release⁶. It was therefore thought prudent to recommend that the rate of water extraction should not exceed the natural recharge rate, as lowering of the water level would reduce the pressure at depth.

S.J. FREETH*

*Geological Hazards Research Unit,
University College of Swansea,
Swansea SA12 8PP, UK*

G.W. KLING

Woods Hole, USA

M. KUSAKABE

Okayama University

J. MALEY

ORSTOM, Montpellier

F.M. TCHOUA

University of Yaoundé

K. TIETZE

BGR, Hannover

*Convenor and Chairman of the Inaugural Meeting of the International Working Group on Crater Lakes.

Mitogen signal

SIR—The recent reports by Xu *et al.*¹ and Buchberg *et al.*² that the proteins encoded in the human and mouse neurofibromatosis type 1 (NF-1) genes are structurally related to the mammalian GTPase-activating (GAP) and yeast IRA1/2 proteins immediately suggest a biochemical explanation for the aberrant proliferation of melanocytes and Schwann cells that is one of the hallmarks of NF-1. This explanation, not advanced in either report, derives directly from the demonstrated ability of cyclic AMP to trigger Schwann cell and melanocyte division.

The following features of signal transduction and Schwann cell and melanocyte biology apply: (1) GAP (in mammalian cells) and IRA1 and IRA2 (in *Saccharomyces cerevisiae*) bind to small G proteins (such as Ras and Ras-related proteins), stimulate the low intrinsic GTPase activity of these proteins, and thereby inactivate them³. (2) A critical target of both Ras in yeast and the G₁₂ proteins of higher eukaryotes is adenylate cyclase. When complexed with GTP, these G proteins activate cyclase, and thereby stimulate the synthesis of intracellular cAMP⁴. (3) Cyclic AMP is a strong but cell-specific mitogen for Schwann cells, melanocytes and certain other mammalian cells^{5,6}. A key component of cAMP mitogenesis in these cells is the cAMP-dependent induction of receptors for polypeptide growth factors such as platelet-derived growth factor⁷. These facts suggest that, just as IRA1 and IRA2 are negative regulators of the G-protein-cAMP pathway in yeast, so NF-1 might, in normal circumstances, stimulate the low intrinsic GTPase activity of a

specialized small G protein in Schwann cells and melanocytes, inactivate this G protein, and inhibit the stimulation of cyclase. Thus, even partial loss-of-function mutations in the NF-1 gene would result in elevation of intracellular cAMP, upregulated expression of polypeptide growth-factor receptors, and increased Schwann-cell and melanocyte proliferation.

The problem with this hypothesis is that the only organism in which an IRA-like protein is known to regulate the cyclase-stimulating activity of a small G protein is *S. cerevisiae*. That the NF-1 protein might be a negative regulator of cAMP metabolism in certain mammalian cells is nonetheless consistent with the autosomal dominant genetics of NF-1 transmission, the known interaction of IRA1/2, Ras and cyclase proteins in yeast, the fact that the NF-1 protein is more similar in amino-acid sequence to IRA1 and IRA2 than to GAP itself, and the demonstration that elevation of cAMP strongly stimulates Schwann cell and melanocyte proliferation.

This hypothesis also makes two predictions that are relatively easy to test: both intracellular cAMP and the expression of cAMP-dependent genes⁸ should be elevated in NF-1 Schwann cells and melanocytes; and expression of the NF-1 phenotype by these cells should be growth-factor-dependent.

GREG LEMKE

*The Salk Institute, PO Box 85800,
San Diego, California 92186-5800, USA*

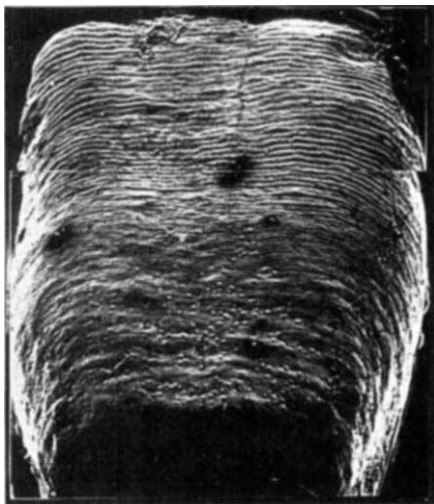
1. Freeth, S. J. & Kay, R. L. F. *Nature* **328**, 299–300 (1987).
2. Kling, G. W. *et al.* *Science* **236**, 169–175 (1987).
3. Sigvaldason, G. E. *J. volcan. Geotherm. Res.* **39**, 97 (1989).
4. Kusakabe, M. *IWGCL Newsletter* **1**, 1 (1989).
5. Sigurdsson, H. *et al.* *J. volcan. Geotherm. Res.* **31**, 1 (1987).
6. Tietze, K. *BGR Report 100 470 1–34* (1987).

1. Xu, G. *et al.* *Cell* **62**, 599 (1990).
2. Buchberg, A. M. *et al.* *Nature* **347**, 291 (1990).
3. Hall, A. *Science* **249**, 635 (1990).
4. Toda, T. *et al.* *Cell* **40**, 27 (1985).
5. Raff, M. C. *et al.* *Cell* **15**, 813 (1978).
6. Dumont, J. E. *et al.* *Trends biochem. Sci.* **14**, 67 (1989).
7. Weinmaster, G. & Lemke, G. *EMBO J.* **9**, 915 (1990).
8. Monuki, E. S. *et al.* *Science* **249**, 1300 (1990).

Dental caution

SIR—The study of Plio-Pleistocene hominid dentition^{1,2} has challenged the notion of a prolonged maturation period throughout human evolution³. New information indicates that some of these conclusions^{1,2,3} are premature and that additional research is needed to clarify the issue.

Bromage and Dean¹ counted enamel surface manifestations (perikymata) of internal bands (striae of Retzius) on nine australopithecine incisor teeth, reporting a range of 57–135 perikymata. They used these data to calculate crown-formation times and the age at death of six fossil specimens with just-completed permanent



The complete dentition of this individual was figured in ref. 11. Here, a montage at 9× of an upper incisor illustrates the perikymata on the labial surface of the tooth. By the method of Bromage and Dean¹, incisor perikymata counts give a crown formation time of 2.2 ± 0.1 to 3.7 ± 0.2 years, or an age at death between the late second and early fourth year in the presence of recently erupted M1s. Estimates for the time of root formation are added to crown formation times for age at death, and here follow previously published estimates for human incisors with 1–3 mm root^{10,12}.

incisors. Compared with a sample of 10 modern human incisors, with a range of 165–202 perikymata, the authors concluded that growth periods for Plio-Pleistocene hominids were similar to the modern great apes. These 10 human teeth, not further identified, remain the only published sample of human incisor perikymata counts, and are the accepted standard for taxonomic interpretation³.

To investigate further the pattern of perikymata in modern humans, we sputter-coated a sample of 12 modern human incisor teeth with gold palladium and viewed them with a scanning electron microscope at 20, 40 and 80× magnification at 6, 12 and 20 KeV, and recorded the labial surfaces by montage images on polaroid film. All teeth had complete crowns with minimal root and had perikymata observable over the entire

crown surface at low binocular magnification. Ten specimens are from the archaeological site of Hasanlu (Iran), about 3000 BC, and two are from the Island Field site in Delaware, about AD 800.

In our sample, perikymata counts range from 75 to 157 (s.d. of 1–12 for individual teeth; overall mean 116; median 118; s.d. 25). These results overlap, within measurement error, the fossil sample, the original modern human range¹, and are similar to those of an unpublished survey by A.-M. Bacon on 23 additional human incisors, with a range of 111–179 perikymata.

Furthermore, perikymata are variable in expression, often missing from the cervical aspect of maxillary incisors. The use of these surface features alone to infer incremental patterns of underlying enamel is questionable. Three teeth in the present sample are from a single individual, with perikymata counts of 75 ± 7 (upper RI1), 128 ± 4 (upper RI2), and 157 ± 12 (lower RI1) (see figure). It is possible that either abrasive phenomena have differentially affected the right maxillary I1, or more internal striae exist than are reaching the surface of the tooth. If incomplete perikymata counts are a feature of enamel and sectioning of teeth is required to count internal incremental lines for accurate estimation of total crown formation time, then enamel surface counts on fossil specimens are inadequate estimators of age at death, and previous reports using these as databases need to be reconsidered^{1,2,5–7}.

Chronological ages derived from perikymata counts have been used to calibrate other dental events, particularly first molar eruption^{2,7}. The incisor in the figure comes from a modern human whose dental sequence parallels that described for the *Australopithecus africanus* mandibular specimen, Sts 24 (refs 4, 8), and similar to Taung⁹ and the *A. afarensis* LH 2 (ref. 10), including an erupted first molar¹¹. This modern human is experiencing first molar eruption between three and four years of age, an abbreviated growth period relative to modern man, similar to the modern great apes¹ and to that presented for *Australopithecus*².

The examination of additional modern

human incisors has broadened the range of variation for perikymata numbers, and results in the significant overlapping with that reported for fossil juveniles^{1,2}. Initial conclusions of ape-like growth periods for Plio-Pleistocene hominids based on ages at death derived from perikymata counts are premature, and their further use to substantiate major taxonomic conclusions^{2,7} should be held in abeyance.

ALAN MANN
JANET MONGE
MICHELLE LAMPL

Department of Anthropology,
University Museum,
University of Pennsylvania,
Philadelphia,
Pennsylvania 19104, USA

Third eye

SIR—Professor May, in his News and Views article¹ on the paper by Daugherty *et al.*² concerning the taxonomic classification of the tuatara (*Sphenodon*), says, as an aside on the animal's third eye, "most other vertebrates retain this organ in vestigial form as the pineal gland".

Vestigial organs have tended to disappear with advances in medical knowledge. As far back as 1931 Sir James Purves Stewart, writing in the 7th edition of his famous work *The Diagnosis of Nervous Diseases*, states "The pineal gland was formerly regarded as a 'vestigial organ', a portion of which in some reptiles develops into a rudimentary para-pineal or parietal eye. Modern observations, however, show that so far as being a mere vestigial relic, it is a glandular structure whose activity is necessary for normal metabolism in early life and that prior to the age of puberty the pineal gland is an important organ of internal secretion."

The modern observations to which Sir James refers are by Professor A. Munzer writing in 1911 and the extensive historical review by Dr L. J. Kidd in 1913 (refs. 3, 4). Sir Cecil Wakeley, past president of the Royal College of Surgeons of England, in his *Faber Medical Dictionary* (1975) gives the pineal body as a structure of unknown function. The best medical report for a layman is to be found in the 37th edition of *Gray's Anatomy*, where it is clearly stated that the pineal is a powerful regulator of enzymes associated with melatonin synthesis and the output of gonadotrophins by the adenohypophysis.

F.W. COUSINS

43 Emanuel House,
18 Rochester Row,
London SW1P 1BS, UK

1. Bromage, T.G. & Dean, M.C. *Nature* **317**, 525–527 (1985).
2. Beynon, A.D. & Dean, M.C. *Nature* **335**, 509–514 (1988).
3. Beynon, A.D. & Wood, B.A. *Nature* **326**, 493–496 (1987).
4. Mann, A.E. *Some Paleodemographic Aspects of the South African Australopithecines* (University of Pennsylvania, Philadelphia, 1975).
5. Bromage, T.G. *J. hum. Evol.* **16**, 257–272 (1987).
6. Dean, M.C., Stringer, C.B. & Bromage, T.G. *Am. J. phys. Anthropol.* **70**, 301–309 (1986).
7. Dean, M.C. *Folia Primat.* **53**, 160–176 (1989).
8. Dean, M.C. *Am. J. phys. Anthropol.* **67**, 251–257 (1985).
9. Conroy, G.C. & Vannier, M.W. *Nature* **329**, 625–627 (1987).
10. Smith, B.H. *Nature* **323**, 327–330 (1986).
11. Mann, A.E. *Nature* **333**, 123 (1988).
12. Moorrees, C.F.A. *et al. J. dent. Res.* **42**, 1490–1502 (1963).

1. May, R.M. *Nature* **347**, 129–130 (1990).
2. Daugherty, C.H., Cree, A., Hay, J. M. & Thompson, M. B. *Nature* **347**, 177–179 (1990).
3. Munzer, A. *Berliner Klinische Wochenschrift* **48**, 1669–1676 (1911).
4. Kidd, L.J. *Rev. Neurol. Psych.* **11**, 1–24; 55–75 (1913).

Evaluating ozone depletion potentials

SIR—Fisher and co-workers attack the difficult and important issues of ozone depletion potentials¹ and greenhouse warming effects² of hydrohalocarbons proposed as environmentally preferable substitutes for the fully halogenated chlorofluorocarbons (CFCs). A significant fraction of hydrohalocarbons are destroyed in the troposphere, so that they deliver substantially less chlorine to the stratosphere per kilogram used and have shorter atmospheric lifetimes than the chemically inert CFCs. The chlorine loading potentials (the amount of chlorine delivered to the stratosphere), the ozone depletion potentials (the change in the ozone column relative to that caused by the reference molecule CFC-11) and lifetimes for other chlorofluorocarbons³ are given in the table. The shorter lifetimes of the hydrohalocarbons imply that these compounds will be removed more rapidly from the atmosphere than CFCs if production ceases, alleviating concern about possible errors in analysis of their probable environmental impacts.

Nonetheless, accurate estimation of the ozone depletion potentials (ODPs) of alternative compounds requires the use of numerical models with a credible simulation of the physical and chemical processes whereby stratospheric chlorine destroys ozone. Fisher *et al.*¹ based their estimates on gas-phase photochemical-model calculations, which predict a modest maximum ozone depletion of perhaps 5% in the next century based on present chlorofluorocarbon release rates, in sharp contrast to observations of 50% ozone losses during the Antarctic spring⁴ and 3–10% in mid-latitudes of the Southern Hemisphere⁵. The role of chlorofluorocarbons in polar ozone depletion has been established by extensive field measurements^{6,9}, yet the implications of polar ozone loss for the calculation of ODPs were not considered by Fisher *et al.*¹.

Gas-phase photochemical-model calculations predict ozone destruction mainly at middle and low latitudes at altitudes from about 25 to 50 km. Because most of the total column is located at lower levels near 20 km, the corresponding column depletion is at most a few per cent for the present atmosphere. The ODPs derived using gas-phase chemical models are

fundamentally related to the chlorine delivered to the stratosphere and subsequently released by the proposed alternative compounds in the ozone destruction region compared to that of CFC-11. This molecule is rapidly broken down in the stratosphere to release its chlorine, whereas some of the other compounds (for example, HCF₂Cl or HCFC-22) have much longer stratospheric lifetimes, and hence release proportionately less chlorine in the ozone-depletion regions predicted by gas-phase models.

But the Antarctic ozone hole is driven by heterogeneous chemistry on the surfaces of polar stratospheric clouds occurring near 10–22 km (refs 5–11), a region where some compounds release far more chlorine. Similar chemistry is found in the Arctic⁷. Past records indicate that only in April are polar stratospheric clouds not present in the atmosphere, so that ozone loss arising from their interaction with chlorine species could occur over much of the year at one pole or the other^{8,12}. Air enters the stratosphere largely in the tropics and descends in polar regions. Thus, air reaching the 20-km region above Antarctica has resided longer in the stratosphere and risen higher than that in mid-latitudes, allowing greater release of ozone-destroying chlorine. Observations of long-lived atmospheric tracers suggest rapid downward transport in the Antarctic stratosphere and far lower abundances of N₂O, CH₄, CFC-11, CFC-12 and, by extension, HCFC-22, than predicted by models^{13,14} such as those used by Fisher *et al.*¹. These observations suggest that HCFC-22 will be largely dissociated by the time it arrives in the region of the Antarctic ozone hole, implying that the ozone destruction there exceeds the ODP and may approach the chlorine loading potential for that gas (see table). The impact of polar processes is greater the larger the difference between the ODP and the chlorine loading potential of the alternative compound (see table).

In the Antarctic atmosphere, reactions involving bromine contribute about 20% of the total ozone loss^{5,9,11}. Yet Fisher *et al.*¹ note that bromine reactions do not contribute to their ODP calculations. This is because the calculated ozone destruction occurs near 25–50 km, where bromine

plays little part, in contrast to that actually observed to take place in polar regions.

If a globally averaged ODP is required for policy formation, such estimates must be based on the physical insight acquired from observations¹⁵ with a view to evaluating errors in numerical models; the implications of polar ozone depletion for global ozone loss cannot be overlooked. As noted by Fisher *et al.*¹, uncertainties in the rate of tropospheric loss of the alternative compounds and hence in the chlorine-loading potentials must also be considered when evaluating the overall uncertainties in ODPs. Fisher *et al.*¹ note the possible role of dilution of the Antarctic ozone hole and transport from the Antarctic to lower latitudes (sometimes called the chemical processor mechanism¹²). The ozone depletion observed in mid-latitudes of the Southern Hemisphere greatly exceeds that explicable by gas-phase photochemistry⁵, suggesting that much if not all of the Southern Hemisphere may be influenced by Antarctic ozone depletion. This, in turn, suggests that the Southern Hemisphere ODPs for HCFC-22 and HCFC-142b exceed the estimates by Fisher *et al.*¹ by at least a factor of two (see table). In the Northern Hemisphere, smaller absolute ozone decreases are observed⁵ but these are still several times greater than predicted by gas-phase photochemical models and are probably related at least in part to polar chemistry, again implying increases in ODPs for such species.

The physical processes described here indicate that the globally averaged ODP estimates provided by Fisher *et al.*¹ are probably lower limits on the actual ozone depletion by any compound for which the chlorine-loading potential substantially exceeds its ODP. The chlorine-loading potentials represent upper limits that are probably approached in polar regions, with substantial effects spreading to lower latitudes.

SUSAN SOLOMON

ADRIAN TUCK

NOAA Aeronomy Laboratory,
Boulder, Colorado 80303, USA

FISHER *ET AL.* REPLY:—Our ability to forecast future levels of global ozone is limited by current understanding of polar chemical processes. Several chemical mechanisms that are not part of the homogeneous chlorine chemical theory of Molina and Rowland¹⁶ may be unique to the polar region. We agree with Solomon and Tuck that understanding polar processes is central to discussions of policy involving stratospheric ozone.

Solomon and Tuck question whether we have underestimated the ODPs of some chemicals by discussing only global results from homogeneous photochemical models¹. We did not discuss relative effects on high-latitude stratospheric

GLOBAL AND HIGH-LATITUDE ODPs, LIFETIMES AND CHLORINE LOADING POTENTIALS

	Global ODP	High-latitude ODP (Θ,S)	High-latitude ODP (Θ,S)/global ODP	Lifetime (yr)	Chlorine loading potential
HCFC-22	0.03–0.07	0.04–0.09	1.13–1.4	15–25	0.15–0.20
HCFC-123	0.013–0.027	0.019–0.029	1.02–1.1	1.5–2.4	0.017–0.022
HCFC-124	0.013–0.03	0.014–0.04	1.08–1.3	5–10	0.04–0.05
HCFC-141b	0.07–0.14	0.09–0.15	1.03–1.1	6–11	0.10–0.15
HCFC-142b	0.035–0.077	0.04–0.10	1.14–1.4	15–25	0.15–0.20
CFC-12	0.88–0.92	0.90–1.0	1.04–1.3	95–120	1.3–1.9

ODPs from two-dimensional atmospheric models of gas phase reactions.

ozone because of space constraints. We focused on the calculated relative effects as traditionally reported and as included in the Montreal Protocol. Because hydrohalocarbons are being considered as substitutes for CFCs it is necessary to examine the relative effects of each chemical on stratospheric ozone without losing sight of the fact that it is the absolute change in ozone that is of concern.

We discussed implications of polar processes on modelling results in detail in the World Meteorological Organization report³. It was that report that was to provide the basis for policy decisions, rather than our *Nature* paper.

We will discuss each of three points made by Solomon and Tuck about the effect of polar ozone processes on the ODPs we reported: (1) the impact of heterogeneous chemistry; (2) the chlorine delivered to the polar regions; and (3) the effects of bromine chemistry on ODPs.

(1) Assuming that models calculate the 'correct' distribution of dissociated radicals released from each of the halocarbons, the question is how sensitive the calculated (relative) ODPs are to chemical processes (both gas-phase and heterogeneous). We agree that inclusion of heterogeneous processes increase the absolute impact on ozone of the CFCs and hydrohalocarbons. We believe, however, that the local ODP (Θ, S) (where Θ and S designate the latitude and season) calculated by the two-dimensional models is, to the first order, independent of chemistry because ODP (Θ, S) is a relative quantity. So if at any location models underestimate the effect on ozone of chlorine released from CFC-11, they do the same for chlorine released from other halocarbons.

ODPs are reported based on changes in the global amount of ozone. To arrive at a proper estimate for global values, relative effects need to be averaged with respect to latitude and season. If the global ozone change were dominated by heterogeneous processes during springtime at high latitudes, the global ODP would be equal to the local ODP at those latitudes during that season. Our calculations showed that the calculated high-latitude ODPs (see table) are within 40% of the global ODP we reported. Because the ODPs for the substitutes are typically less than 0.1, even in the extreme case when a 40% margin applies, the effects from substitutes would therefore be considerably less than those from CFCs.

(2) How reliable are the calculated distributions of reactive chlorine released by various halocarbons? More specifically, do models underestimate the amount of chlorine released by certain halocarbons (such as HCFC-22) to the polar stratosphere? Although the data needed to settle this question (measurement of polar profiles) are incomplete, model results indicate

that the downwelling effect suggested by Solomon and Tuck should have only a minor effect on local ODPs.

The argument is as follows. Measured CFC-11 and CFC-12 concentration profiles over Antarctica¹⁴ exhibit about 5-km shifts downward compared to those predicted by models suggesting that models may underestimate the downwelling in this region, as well as the amount of dissociated chlorine. The ODP (Θ, S) for HCFC-22 can then be estimated by shifting the profiles for CFC-11 and HCFC-22 by the same amount. The new ODP (Θ, S) evaluated using the shifted profiles would not be significantly larger than we have calculated because the stratospheric scale height for HCFC-22 is greater than the scale height for CFC-11.

The distributions of the source gases and chlorine radicals in the stratosphere depend only on transport and gas-phase chemistry (specifically photolysis, reactions with $O(^1D)$ and/or OH). Heterogeneous reactions are not important for degradation of these compounds.

The WMO report included results from a series of sensitivity tests using the Atmospheric and Environmental Research two-dimensional model in which transport was modified to increase chlorine delivered to the polar region. The calculated local ODP for HCFC-22 was around 0.1 for most cases, although it could be made as large as 0.15 in one case. The simulated global ozone distribution for this one case did not, however, resemble measured values.

(3) Solomon and Tuck correctly point out that bromine chemistry can be important in polar ozone chemistry, but their argument fails to show how it would affect the relative measure of ODP. The reason that bromine chemistry has little effect on ODP is the following. First, the rate-limiting step for the bromine-chlorine cycle is the reaction of BrO with ClO so the rate is determined by the product of the concentrations of these two compounds. Second, because the bromine source gases dissociate at lower altitudes¹⁷ than do the chlorine source gases, the integrated reaction rate is proportional to the spatial distribution of dissociated chlorine. Thus the calculated ozone decrease has the same functional dependency on dissociated chlorine whether one is considering the $ClO-BrO$ process or the processes catalysed by chlorine alone.

Whereas it is ideal, as Solomon and Tuck advocate, to derive ODPs based on physical insight, there is no model-independent methodology for doing so. Solomon and Tuck correctly point out that polar processes may dominate changes in global ozone and should be included in ODP values used for policy making. But, our analysis, summarized above and included in the WMO report³, shows that ODP values may not be as sensi-

tive as speculated by Solomon and Tuck, principally because ODP is a relative measure.

Despite major advances in the qualitative understanding of fundamental processes that control polar chemistry, we are still limited by our inability to quantify key parameters. As a result, models of polar processes cannot yet predict the complex interactions, let alone the role of halocarbons. These investigations require time which is not compatible with immediate policy-making needs. Nevertheless, it is important to understand polar processes and to resolve the questions related to heterogeneous chemistry and chlorine delivery.

DONALD A. FISHER

Scientific Computing Division,
E. I. du Pont de Nemours and Co.
Experimental Station,
Building 320,
PO Box 80320, Wilmington,
Delaware 19880-0320, USA

MALCOLM KO

Atmospheric Environmental Research
Inc.,
840 Memorial Drive,
Cambridge,
Massachusetts 02139, USA

DONALD WUEBBLES

Lawrence Livermore National
Laboratory,
California 94550, USA

IVAR ISAKSEN

Institute of Geophysics,
0315 Blindern,
Oslo 3, Norway

1. Fisher, D.A. *et al.* *Nature* **344**, 508–512 (1990).
2. Fisher, D.A. *et al.* *Nature* **344**, 513–516 (1990).
3. World Meteorological Organization *Scientific Assessment of Stratospheric Ozone* No. 20, Vol. II, AFEAS Rep. (WMO, Geneva, 1990).
4. Farman, J.C., Gardiner, B.G. & Shanklin, J.D. *Nature* **315**, 207 (1985).
5. World Meteorological Organization *Scientific Assessment of Stratospheric Ozone* Rep. No. 20, Vol. I (WMO, Geneva, 1990).
6. Tuck, A.F., Watson, R.T. & Toon, O.B. (eds) *J. Geophys. Res.* **94**, Nos 9, 14 (1989).
7. Turco, R.P., Plumb, R.A. & Condon, E.P. (eds) *Geophys. Res. Lett.* **17**, no. 4 (1990).
8. Anderson, J.G. *et al.* *J. geophys. Res.* **94**, 11480 (1989).
9. de Zafra, R.L. *et al.* *Nature* **328**, 408–411 (1987).
10. Solomon, S. *et al.* *Nature* **321**, 755–758 (1986).
11. McElroy, M.B., Salawitch, R.J., Wofsy, S.C. & Logan, J.A. *Nature* **321**, 759–762 (1986).
12. Tuck, A.F. *Geophys. Res.* **94**, 11687–11737 (1989).
13. Podolske, J.R., Loewenstein, M., Strahan, S.E. & Chan, K.R. *J. geophys. Res.* **94**, 16767–16772 (1989).
14. Heidt, L.E. *et al.* *J. geophys. Res.* **94**, 11599–11611 (1989).
15. Anderson, P.W. *Phys. Today* **43**, 9 (1990).
16. Molina, M.J. & Rowland, F.S. *Nature* **249**, 810–814 (1974).
17. World Meteorological Organization *Atmospheric Ozone 1985*, Rep. No. 6 (WMO, Geneva, 1986).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

Heroes of the revolution

Peter Goodfellow

Genome. By Jeremy E. Bishop and Michael Waldholz. *Simon and Schuster*. 1990. Pp. 352. \$22.95, £16.95.

IN 1959 a vicious murder took place in Holcomb, a small town in Kansas. The novelist Truman Capote became obsessed with this murder; he became so obsessed that he went to live in Holcomb and interviewed everyone he could find who was prepared to comment on the murderers or their victims. Eventually he reported the results of his research and cogitations in the book *In Cold Blood*; this book is a blend of fact and opinion, journalism and history. Using this model, investigative journalists and inquisitive novelists have sought to turn today's news into tomorrow's books. Molecular genetics and molecular biology are newsworthy: genetic engineering and test-tube babies are part of the normal lexicon of the headline writer. The book-of-the-news has followed the headlines. My bookshelf is burdened with *Natural Obsessions*, *Hidden Frontiers* and many other books describing the endeavours, disappointments and breakthroughs of scientists. *Genome* is in this tradition: a blend of anecdote, biography and scientific explanation. The anecdotes are endearing, the biography is inspiring and the science is accurate. It is easy to read and only suffers in comparison to my usual bed-time reading because it lacks sex and violence.

In 1986, the journalist B. Bryson went to Kansas and attempted to investigate the investigation that Capote had completed 20 years previously (*The Lost Continent*; Secker and Warburg, 1989). The good people of Holcomb did not want to know; they had been offended by the picture of their world created by a novelist and refused to collaborate with a journalist. Herein lies a problem, Capote in search of a higher truth was presumably unconcerned about offending his erstwhile hosts. Bishop and Waldorf, the authors of *Genome*, are science journalists working for the *Wall Street Journal*. If they offend their interviewees, will they will be welcome back? Does this colour their descriptions?

Another problem with the faction approach is that it depends critically on who is interviewed. There are those who are called but refuse to answer and there are those who are never called. I would deduce that the budget of the *Wall Street Journal* is insufficient to allow their science journalists to visit Europe and that explains the insular view presented in this book.

Genome bills itself as "The most astonishing scientific adventure of our time —

the attempt to map all the genes in the human body". This hype is misplaced. The book is not concerned with gargantuan struggles between the NIH and the American Department of Energy, nor is it concerned with the relative merits of YACs, cosmids and P1 vectors. In short it has little to do with the Genome Project, instead it describes the impact of molecular biology on medical genetics. The authors rightly identify the restriction fragment-length polymorphisms (RFLPs) as the point of revolution. Armed with an unlimited supply of genetic markers it is possible to locate the chromosomal position of disease loci and this is the first step in the dance that leads to gene isolation, definition and, we hope, better patient care.

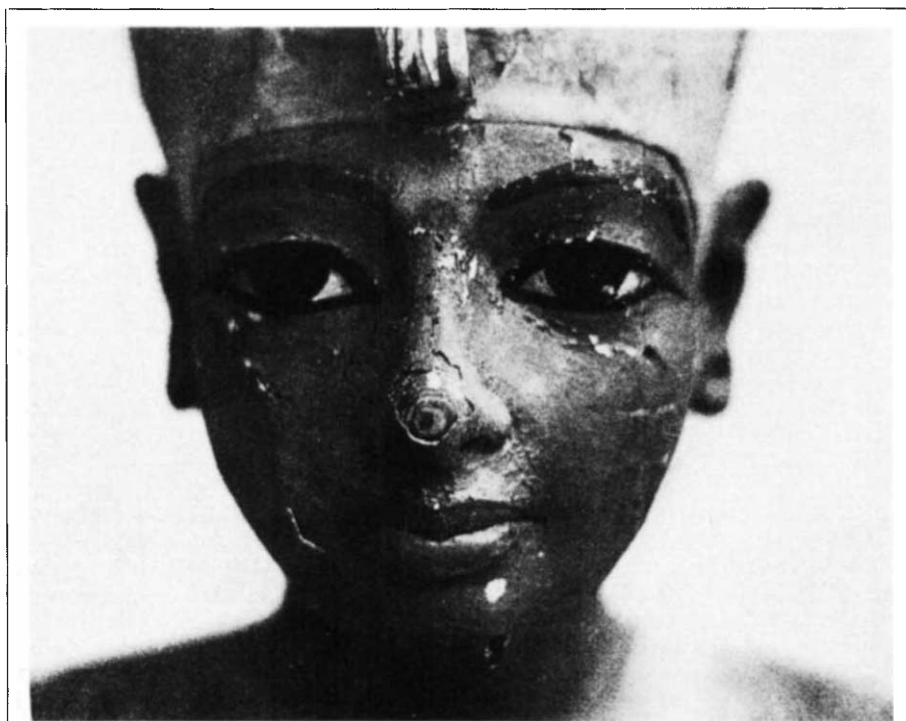
The central figure in the book is David Botstein and the central revelation occurred in the Wasatch mountains of Utah. In the idyllic surroundings of a ski resort, Botstein realized that RFLPs could be used as markers in human genetics. Back in the more prosaic environment of MIT he recruited brilliant young investigators to exploit this revelation. *Genome*

conjures the image of people in-the-know spreading the word RiFLiP, which is apparently how geneticists recognize each other. Of course, Botstein was a major catalyst, but, from my biased view point, this analysis ignores the contributions made by A. Jeffreys, E. Solomon, W. F. Bodmer and many others. It also ignores our cultural heritage as geneticists: RFLP analysis is based on the techniques developed by R. A. Fisher, J. B. S. Haldane, S. Wright, J. H. Renwick, N. Morton, and again, many others. The only chapter in the book that attempts scholarly appraisal of the past derives from interviews with V. McKusick.

After setting the stage with Botstein, the authors follow his disciples as they apply the new methods to different diseases. Again, I do not always agree with the chosen emphasis, but the stories are inspiring. Great progress has been made in the study of muscular dystrophy, cystic fibrosis and several forms of cancer. I am less convinced that progress has been made in studying alcoholism and psychiatric disease.

The strengths of this book should be balanced against its deficiencies. It is very clearly written, easy to follow and does not duck explaining the scientific issues to a lay audience. I shall lend my copy to my parents. □

Peter Goodfellow is at the Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.



The head of a painted wooden mannequin of Tutankhamun, thought to have been used to hang the king's robes or jewellery. *The Complete Tutankhamun: The King, The Tomb, The Royal Treasure* by Nicholas Reeves describes the region of the boy king and the story of Howard Carter and Lord Caernarvon's quest for his tomb in the Valley of the Kings. Published by Thames and Hudson, price is £15.95. □

Species talk

R.H.L. Disney

The Natural History of Blackflies. By Roger W. Crosskey. Wiley: 1990. Pp. 711. £75, \$172.50.

THIS is a landmark publication on a family of flies of enormous medical importance, some veterinary importance and of interest to freshwater biologists. It provides a much needed critical synthesis of a rapidly expanding literature. It is accessible to the nonspecialist and is written with evident enthusiasm. The author's reputation for attention to detail is maintained, except for an uncharacteristic failure to list in the bibliography all the references cited in the text.

Crosskey is primarily a taxonomist of distinction. He has tended to shy away from theoretical discussions and disputes and to avoid formal phylogenetic analysis. His attitude is reflected in his comment that "it is safe to assume that users of this book are unlikely to be much interested in taxonomic theory — or expect to find it here". Although this is, perhaps, a defensible assumption with respect to classification, it is an unfortunate comment when it comes to species recognition problems.

Crosskey highlights the important discovery that many blackfly species are 'complexes'. Almost every 'morpho-species' intensively investigated has turned out to be a complex in terms of cytotaxonomy. Studies on enzymes support these findings. The central theoretical question is with regard to the status of these cytotypes. Are we dealing with sibling species, polymorphism, polytypic species or a mixture of all three? The resort to cross-breeding experiments in the laboratory, which have proved invaluable with mosquitoes, has been largely unavailable with blackflies, which have proved difficult or impossible to culture in the laboratory.

Crosskey not only makes it clear that a lack of gene flow between 'cytotypes' is the criterion for recognizing that a cytotype is a species, but in blackflies it has to be inferred rather than demonstrated. The basic grounds for such an inference are stated to be (1) the chromosomal inversion, or whatever, which defines the cytotype is only found in the homozygous state; (2) there are differences in ecology, behaviour and/or distribution with respect to other 'cytotypes'.

Most cytotypes have been established on relatively small samples of larvae. As larger samples are examined the chances of finding a heterozygote must increase. With the intensively studied vectors of Onchocerciasis, the *Simulium damnosum* complex, this is precisely what has happened. Crosskey summarizes the present

position thus: "Natural hybrids occur between *S. damnosum* complex siblings, and Boake and Mosha... recently identified (by chromosomal features) a cross between the distantly related *S. sirbanum* and *S. sanctipauli*."

We must ask, therefore, why in the face of this evidence of invalidation of criterion 1 we do not reject the hypothesis of sibling species? The answer offered is that criterion 2 still stands. There are correlations between cytotypes and ecological differences. This argument certainly carries weight. But it does not settle the matter. On the contrary, it assumes that ecological and ethological uniformity is a property of species. Ernst Mayr has aptly warned (*Populations, Species, and Evolution*; Harvard University Press, 1970): "The more closely a species is studied, the more likely it is that some evidence will be

found for ecological polymorphism or gradual ecological variation. The variation is still largely ignored by ecologists, most of whom discuss the ecological requirements of species in a strictly typological manner."

With the data on *Simulium damnosum* it is entirely defensible to challenge the current orthodoxy, which equates cytotypes with sibling species. The work of Dunbar and others suggests that these 'cytotypes' fall into about six groupings. Perhaps the latter are the real biological species and the cytotypes are evidence of polymorphism and polytypy within them. Clearly blackflies are an excellent group for exploring these fundamental issues. □

R. H. L. Disney is in the Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.

Fluke success?

F. E. G. Cox

Modern Parasite Biology. Cellular, Immunological and Molecular Aspects. Edited by David J. Wyler. Freeman: 1990. Pp. 428. Hbk £47.95, \$59.95; Pbk £32.95, \$42.95.

THE gap between textbooks of parasitology and the research literature has never been greater and over the last two years there have been a number of attempts to bridge this gulf. All have been multi-author texts and some have been rather specialized books. These include *Vaccination Strategies of Tropical Diseases*, edited by F. Y. Liew (CRC), *Immunity and Pathology*, edited by J. M. Behnke (Taylor and Francis) and *Parasite Communities*, edited by G. Esch, A. Bush and J. Aho (Taylor and Francis). Others have attempted a more general coverage of the field such as *New Strategies in Parasitology*, edited by K. P. W. J. McAdam (Livingstone) and *The Biology of Parasitism*, edited by P. T. Englund and A. Sher (Liss). David Wyler's book belongs to the latter category and is based on two assumptions; that modern parasitology is largely concerned with immunology and cell and molecular biology, and that the parasites of most interest are the major pathogens of humans. The book is aimed at a wide audience with interests in these areas.

For convenience this book is divided into three parts. The first, cell biology, contains contributions on *Plasmodium*, African trypanosomes, American trypanosomes, *Leishmania*, *Toxoplasma*, *Entamoeba* and *Schistosoma*. The second, immunology, is concerned with infections caused by the same organisms except for amoebiasis, which is covered briefly in the first part, plus the addition of lymphatic filariasis. The third, molecular biology,

comprises chapters on *Plasmodium*, trypanosomes (essentially African), *Leishmania*, and schistosomes and filarial worms in a single chapter. In all, there are 19 chapters written by 26 acknowledged authorities in their fields.

There is, therefore, a potential feast here for anybody interested in these rapidly developing areas of parasitology, but it is, unfortunately, both stale and somewhat indigestible. Each chapter contains suggestions for additional reading but the lists are limited and dated. For example, two chapters on the cell biology of schistosomes list only five references (one subdivided into parts) of which only one is from the 1980s. There are, however, about 450 references to primary source papers but these are almost entirely from the early 1980s and are not cross-referenced in the chapters themselves. There is also rather too much overlap, thus antigenic variation in trypanosomes is covered three times. The helminth worms receive less attention than they warrant considering the extent of our understanding of the immunology and molecular biology of the nematodes and cestodes.

Overall, this is a worthy book and one that is well worth reading provided that it is realized that modern does not mean recent, that the coverage is limited and that there is nothing in it concerning some of the most exciting discoveries, such as RNA editing and the roles of cytokines, which have transformed the molecular and immunological aspects of parasitology over the last five years. □

F. E. G. Cox is in the Division of Biomolecular Sciences, King's College London, Campden Hill Road, London W8 7AH, UK

■ Just published by Open University Press is *Molecular Parasitology* by J. E. Hyde, which describes how molecular biology has led to more detailed analyses of parasitic systems. It should be of interest to students and senior workers. Price is £45 (hbk), £18.99 (pbk).

In Newton's shadow

Willem Hackmann

Before Newton. The Life and Times of Isaac Barrow. Edited by Mordechai Feingold. Cambridge University Press: 1990. Pp. 380. £35, \$49.50.

When one thinks of Isaac Barrow one immediately thinks of Isaac Newton, and that has been the misfortune of one of the most prominent seventeenth-century mathematicians and men of science. History has not treated him kindly. Barrow's reputation was quickly put in the shade by the successor to his Lucasian chair of mathematics at Cambridge, the incomparable Newton, whom Barrow supported successfully for the post. His reward has been that he has remained a rather shadowy figure which this biography, consisting of a collection of chapters by Barrovian authorities, attempts to illuminate.

According to an earlier biographer, Barrow was known to his contemporaries for his slovenly ways, his excessively long sermons, and his great learning. He is best remembered for his mathematical and optical writings. History always has difficulty in judging a polymath and that is what Barrow was in seventeenth-century terms. Others have judged him an ambitious careerist. Certainly, as this biography shows, ability did not get one far without patronage. Barrow was lucky with his patrons both in and outside Cambridge. He was educated at Charterhouse, Felstead School and Trinity College, Cambridge, where he was successively a pensioner (1643), college fellow (1649), and Master (1673). His first attempt at the Regius professorship of Greek failed for politico-religious reasons (Cambridge was in the throes of the Civil War), but his second attempt (1660) at the restoration of the monarchy was successful. Subsequently he became Gresham professor of geometry in London (1662), thereby supplementing his meagre Cambridge income, the first Lucasian professor of mathematics (1663), Royal Chaplain to Charles II in London (1669), and finally Vice-Chancellor of Cambridge (1675).

Barrow's misfortune was, as Feingold points out in his introductory chapter 'Isaac Barrow: divine, scholar, mathematician', that he was active at the turning point of classical culture in the mid-seven-

teenth century — a learned culture that became virtually obsolete in Barrow's lifetime, when a more narrow scholarly professionalism started to creep in. Past historians of science, too, have largely ignored Barrow in their quest for the 'giants' of the scientific revolution, and his name is linked to no discovery. Thus, Barrow became chiefly remembered as the mentor of Newton, but even this relationship has never been clarified. That Newton was Barrow's pupil at Trinity is a myth. His name does not appear in



Newton's surviving papers, and there is no evidence to suppose that Newton's early mathematical or optical discoveries were due to Barrow. These aspects are discussed in both Feingold's introductory chapter and in John Gascoine's chapter, entitled 'Isaac Barrow's academic milieu: Interregnum and Restoration Cambridge'. Both chapters also give a good insight into mid-seventeenth-century university education at Cambridge.

Insight into the relationship between Barrow and Newton is also somewhat unexpectedly given in the final chapter in this volume in which Feingold describes Barrow's library. As his first biographer, Abraham Hill, wrote, the estate left by Barrow "was books". Of his library of 1,100 volumes, predictably for the period more

than 80 per cent were in Greek or Latin, and the remainder in English, French, Italian, Hebrew and Arabic. The books were almost evenly distributed among Barrow's three main interests: theology, science (about half of these were mathematical and the remainder were devoted to medicine and philosophy) and the humanities. Newton had free access to Barrow's library and this is the reason why he was not compelled, as Feingold suggests, to acquire a large number of books before Barrow's death in 1677. The largest single category of books purchased by Newton before 1677 was alchemy, a subject conspicuously absent from Barrow's library catalogue.

Barrow's optical and mathematical contributions have been very adequately dealt with in Alan

Shapiro's chapter, entitled 'The

Optical Lectures and the foundations of the theory of optical

imagery', and in Michael

Mahony's 'Barrow's mathe-

matics: between ancients

and moderns'. Shapiro

points out that Barrow

completed a major phase

of the Keplerian revolution

in geometrical optics by creating a

mathematical theory of

optical imagery. The

weakness was that he did

not deal with applied

optics and such topics as

the eye and vision, the

telescope and microscope.

Other biographers, notably

Whiteside, have been less

kindly about Barrow's optics,

but he does agree that his most

original contributions were his

method for finding the point of

refraction at a plane interface and

his point construction of the dia-

caustic of a spherical interface. In

mathematics Barrow and Newton were

pursuing quite different paths. Barrow

preferred the analysis of geometrical

configurations; Newton's fluxions (his

form of calculus) clearly lay in the

algebraic tradition.

Barrow emerges from these pages a

scholar in the humanist tradition (VIDE

Anthony Grafton's brief chapter, 'Barrow

as a scholar'), a sincere man with a gentle

temper (see Irène Simon's equally short

chapter, 'The preacher') and, to give the

final word to the editor of this volume, a

man whose integrity is the recurring motif

of his life. □

Willem Hackmann is in the Museum of the History of Science, Broad Street, Oxford OX1 3AZ, UK.

The picture shows Isaac Barrow in 1676 (courtesy of the National Portrait Gallery).

Fossils for beginners

Alec Panchen

Vertebrate Palaeontology: Biology and Evolution. By Michael J. Benton. *Hyman*: 1990. Pp. 377. Hbk £45; pbk £14.95.

A FEW years ago a debate was staged at the annual meeting of British vertebrate palaeontologists to discuss the question "What is vertebrate palaeontology for?" It was not an unqualified success, rapidly degenerating into the usual scientists' exchange about inadequate funding, but the question remains an important one. Should students of vertebrate fossils be concerned solely with an attempt to reconstruct the anatomy and way of life of the extinct creatures that pre-occupy them, thus to produce a reconstructed phylogeny and narrative history of the group, or should they also aim for a conscious input from their studies to evolutionary theory and historiography as well as to history of the Earth's biota?

The traditional response has been to duck the question. For about fifty years the standard text and reference was Romer's *Vertebrate Paleontology* which raised the narrative tradition to a high level. Its successor is Carroll's *Vertebrate Paleontology and Evolution* with a format closely modelled on Romer, but with more latitude given to controversy and more overt reference to primary sources. But the aims are the same, and so they are in Michael Benton's *Vertebrate Palaeontology, Biology and Evolution*, which is intended for a less palaeontologically sophisticated audience than Carroll. Within its limits it is very successful. It is written in a clear and engaging style and is well referenced and generally up-to-date. The illustrations, all line drawings, are mostly skeletal reconstructions of fossils, but with cladograms, evolutionary trees, diagrams of living organisms and palaeogeographical reconstructions as appropriate. All are useful and adequately presented. In the text, Benton opts for a cladistic approach to vertebrate classification, as do most vertebrate palaeontologists, but I rather doubt whether his brief and uncharacteristically opaque account of cladistic methods is adequate as background for their use in the rest of the book. The order of chapters is a compromise between the traditional taxonomic sequence, origin of vertebrates and Agnatha first, human evolution last, and one of geological time, so that post-Devonian fishes come after Triassic reptiles. But it would be wrong to give the impression that the book is just one damned fossil after another. Descriptive sections are interspersed with brief but

lucid accounts of famous fossil communities, extinction events, phylogenetic summaries and so on. Overall it presents an interesting account of the parade of vertebrate history.

But still the narrative tradition is not breached. I can imagine two other kinds of vertebrate palaeontology text which might command a wider, or at least a different audience. The first would concentrate on the methods, not just the practical techniques but also the modes of thought, of vertebrate palaeontologists. Symptomatic would be the illustration, perhaps as photographs, of actual fossil specimens in all their imperfection. It is ironical that for such information one has to turn to books written for a mass audience, notably those on dinosaurs. The second kind of text would try to tackle our question "What is vertebrate palaeontology for?" It would discuss the nature and validity of data and hypotheses from vertebrate palaeontology as contributions to the debate on systematics, the patterns and processes of phylogeny, the causes of mass extinction, the nature and significance of fossil communities and other con-

troversies where the lead has been taken by nonvertebrate palaeontologists.

Benton's book fits neither of these patterns, but nevertheless is an introduction to vertebrate history which might well fire the imagination of students that will go on to study these weightier problems. □

Alec Panchen is Reader in Vertebrate Zoology, Department of Biology, University of Newcastle upon Tyne, NE1 7RU, UK.

■ *Paleontology of Vertebrates* by Jean Chaline describes the evolution of all of the major vertebrate groups from the Agnatha to the hominids, taking a cladistic approach throughout. Also discussed are modes of fossilization, palaeontological techniques, the species concept and the place of palaeontology in evolutionary biology as a whole. Published by Springer, price \$29.50. New out in paperback is *Evolution and the Fossil Record* edited by Keith Allen and Derek Briggs (for review, see *Nature* 340, 684; 1989). The book is concerned with the many aspects of evolution that can only be tackled using evidence gleaned from the fossil record. Published by Smithsonian Institution Press, price \$19.95 (also available in hardback, price \$49.95).

Flights of fancy

Grahame Hardie

Identification of Protein Consensus Sequences. By Alastair Aitken. *Horwood*: 1990. Pp. 167. £31.

THE determination of the complete sequence of a genome represents the end of one phase of research, but also the beginning of a new, more interesting phase. The computer can quickly assign those curious creatures called open reading frames (ORFs) which encode the putative proteins of the organism. Eventually functions may be assigned to these on the basis of similarity with proteins which have been sequenced independently. But an amino-acid sequence remains a bland specimen unless it can be garnished with trimmings: active-site domains, prosthetic group and binding-site motifs, and sites of processing and post-translational modification.

For those with a need to enliven such a dull sequence, this book serves a dual purpose. First, it presents a review of motifs such as potential phosphorylation sites, CAAX boxes and zinc fingers with which speculatively to adorn your sequence. Second, when the journal editor demands some evidence for your flights of fancy, it provides brief practical details, and lists of references, to enable you to obtain it.

The book is particularly valuable for its discussion of post-translational modifications, of which over a hundred have now been described, including phosphoryla-

tion (on serine, threonine, tyrosine or histidine), attachment of various lipids (acetyl, myristoyl, palmitoyl or polyisoprenoids), and carbohydrates (glycosyl groups and glycosyl-phosphatidylinositol membrane anchors). These modifications are often difficult to identify on automated amino-acid sequencers, and N-terminal modifications may preclude their use altogether. The author describes the various technical tricks required to overcome these problems, as well as discussing alternative approaches such as the fast-atom bombardment-mass spectrometer. In other chapters he discusses primary structure motifs such as kinase and protease active sites, metal-binding sites and disulphide bridged repeats. In a book of this length the coverage of such a large topic could not be comprehensive, and the would-be theoretical biochemist armed with his genome sequence and computer terminal will find the book to be a useful introduction rather than a bible. Nevertheless, all of the currently fashionable motifs, such as leucine zippers, GTP-binding sites, EGF-like regions and kringles are discussed. A useful appendix provides a concise listing of the motifs and consensus sequences discussed earlier in the book. Overall, the author has performed a valuable service in bringing together information which is otherwise scattered throughout the biochemical literature, and I feel sure that the book would not remain undisturbed on its library or laboratory shelf for long periods. □

Grahame Hardie is in the Biochemistry Department, The University, Dundee DD1 4HN, UK.

Comparison of terrestrial and marine temperature estimates for the past 135 kyr off southeast Africa: a test for GCM simulations of palaeoclimate

E. Van Campo*, J. C. Duplessy†, W. L. Prell‡, N. Barratt‡ & R. Sabatier§

* Laboratoire de Palynologie, URA 327 du CNRS, Université de Montpellier II, Sciences et Techniques du Languedoc, 34095 Montpellier cedex 5, France

† Centre des Faibles Radioactivités, Laboratoire mixte CNRS-CEA, 91198 Gif-sur-Yvette, France

‡ Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

§ Faculté de Pharmacie, Université de Montpellier I, 34060 Montpellier cedex 1, France

A combined oxygen isotope, planktonic foraminiferal and pollen record from the southwestern Indian Ocean makes possible a comparison between estimated lowland terrestrial temperatures and isotope and faunal sea surface temperatures. The estimates of land and sea mean temperatures over the past 135 kyr are consistent to within 0.5 °C, and in all three records the maximum temperature changes between glacial and interglacial times are about 3–4 °C.

THE relationship between modern and past terrestrial and sea surface temperature (TST and SST) is one of the major uncertainties involved in deriving the climate during the last ice age. The relationship is important to general circulation models (GCMs) of the atmosphere because many of them prescribe SST, which then exerts strong control on the spatial patterns of heat flux to the atmosphere. The question is intriguing because some data for the Last Glacial Maximum indicate that the high variability TST is not consistent with the low variability SST. These discrepancies are obvious in the comparison of CLIMAP^{1,2} SST estimates with nearby TST estimates³.

Webster and Streten⁴ compared TST estimates derived from snow-line lowerings and pollen data on the mountains of New Guinea with SST estimates from the surrounding area¹. Based on modern lapse rates, they concluded that the SST estimates were too high and should be as much as 5 °C lower to be consistent with their interpretation of lowered tropical snow lines, which are clearly documented but less well dated⁵. A Goddard Institute for Space Studies GCM simulation of the Last Glacial Maximum using CLIMAP surface boundary conditions and the insolation value at 18 kyr suggested that CLIMAP tropical SSTs were too high to be consistent with the pollen and snow-line data from low latitudes⁶. Particularly, the annual TSTs during the Last Glacial Maximum were only 3.1 °C lower than present values in East Africa using annual SSTs that were 1 °C cooler. A second simulation, using the same ice-age boundary conditions but SSTs lowered by 2 °C relative to the CLIMAP values, produced a better fit (–5.3 °C) to the East African terrestrial data^{7,8}, pointing to possible uncertainties in the CLIMAP SST reconstructions, the TST reconstructions or the model response. Recently, La Fontaine⁹ has summarized both the palaeoclimate data for equatorial East Africa and a series of palaeoclimate simulations of the Last Glacial Maximum and Holocene using the National Center for Atmospheric Research GCM¹⁰. She confirmed that most of the palaeotemperature estimates come from high-altitude sites and that the geological estimates range from –3 °C to –8 °C with snow-line lowerings

from 500 to 1,000 m. The only quantitative estimate from an East African mid-elevation site¹¹ indicates a temperature decrease of 4 ± 2 °C for the last glacial period. Comparisons with the climate simulations for the Last Glacial Maximum showed that vertical amplification of the temperature changes produced by the models accounted for about half of most of the estimates based on geological observations. The free air lapse rates may not be good estimates of the local lapse rates and processes along mountain sides, however, which have site-specific effects such as convergence, cloudiness, precipitation and winds. Unfortunately, few geological estimates are available from lowlands which might be more directly comparable with SST. Also, SSTs along coasts are subject to more local boundary effects than open ocean temperatures. TST and SST estimates derived from the same record could share the same chronology and regional climate history. Here we report on a combined oxygen isotope, planktonic foraminiferal and pollen record from the southwestern Indian Ocean. This record allows us to estimate lowland temperatures from the pollen record and directly compare them to estimated SST from isotopic ratios and plankton in the same core.

Chronology

Core MD79-254 (12.4 m long) was obtained off Mozambique near the mouth of the Zambezi river (17°53.4' S, 38°40.2' E) at a water depth of 1,934 m (Fig. 1). We (J.C.D.) measured the oxygen and carbon isotopic ratios on benthic foraminifera (*Uvigerina*) and planktonic foraminifera *Globigerinoides ruber* (white) following the method described in ref. 12. All of the standard isotope stages and substages^{13,14} were identified to stage 8 and used to apply the timescale derived by Imbrie *et al.*¹⁵ (Fig. 2). For the past 30,000 years, the $\delta^{18}\text{O}$ can be easily checked by correlation with that of the nearby core MD79-257 (20°24' S, 36°20' E, 1,262 m), which has also been dated by ¹⁴C accelerator mass spectrometry (J.C.D., unpublished data). The age reference points are listed in Table 1. We have linearly interpolated between reference points to estimate the age of each level. The age model indicates that the average accumulation rate is 6.7 cm kyr^{–1} over the past 135 kyr, but varies from ~8.3 cm kyr^{–1} between 0 and 65 kyr to 5.3 cm kyr^{–1} between 65 and 135 kyr. Hence, the resolution of our 10-cm interval is ~1 kyr in the upper part of the record and 2 kyr in the lower part.

Isotope sea surface temperature estimates

The $\delta^{18}\text{O}$ of planktonic foraminifera reflects a complex mixture of fractionation processes, including the mean $\delta^{18}\text{O}$ composition of the world ocean, the local temperature of calcite precipitation, the local evaporation/precipitation ratio and species-specific biological fractionation. To estimate local SST, these effects must be known or constrained. The species-specific fractionation of *G. ruber* is assumed to have been constant during the

TABLE 1 Age and depth of reference points

Depth (cm)	Age (kyr)
140	15.0
320	32.7
370	40.7
440	50.3
490	56.1
520	58.9
540	65.0
660	87.0
750	107.0
860	122.0
880	128.0
910	135.0

Quaternary, because the isotope records for this species are identical to those of *G. sacculifer*, which has a much smaller species-specific fractionation^{16,17}. There is no way to estimate directly the variations of the evaporation/precipitation ratio during the past climate cycle from isotope measurements alone. The modern mean surface water salinity at the core site is close to 35‰, however, and in this salinity range the slope of the relationship linking seawater $\delta^{18}\text{O}$ to salinity is small and close to 0.4 (ref. 18). Salinity estimates derived from the analysis of fauna in the eastern Indian Ocean indicate that salinity changes have been small in the equatorial area far from the influence of the Bay of Bengal¹⁹. As the South Equatorial current is the source of water for the Agulhas current, we estimate that the $\delta^{18}\text{O}$ variation owing to local salinity changes is $<0.1\%$ (equivalent to 0.25‰ change in mean local salinity). Therefore, $\delta^{18}\text{O}$ variations measured in foraminiferal shells only reflect the global variations of the ocean $\delta^{18}\text{O}$ owing to continental ice volume changes and the isotope fractionation between calcium carbonate and water, which depends on the temperature at which foraminifera formed their tests¹¹. Recently Labeyrie *et al.*²⁰ have separated these effects and derived a $\delta^{18}\text{O}$ record that

reflects only the mean seawater $\delta^{18}\text{O}$ variations owing to continental ice volume changes during the past 135 kyr. As the seawater $\delta^{18}\text{O}$ record is common to all isotope records, the temperature variations in the Mozambique Channel can be calculated by subtracting the mean seawater $\delta^{18}\text{O}$ record from the planktonic isotope record of core MD79-254. The difference is then expressed in terms of temperature according to the palaeotemperature equation²¹.

To perform these calculations, the two time series must have the same sample intervals. The mean seawater $\delta^{18}\text{O}$ record has been expressed with a sampling interval of 1 kyr (ref. 22). We therefore interpolated our raw data using a 1-kyr interval, which is close to the actual sampling interval.

We took into account that *Globigerinoides ruber* is not in isotopic equilibrium with ambient water. Detailed measurements made on living plankton and surface sediment shows that the isotopic ratios from *G. ruber* are smaller than the equilibrium value and that this offset may be as high as 0.3–0.6‰^{23,24}. As the present composition of the surface water from the tropical Indian Ocean is close to +0.5‰ (relative to SMOW)¹⁸, the two effects roughly cancel. We therefore applied no correction to estimate the SST from the isotope record. This simplification results only in a constant offset and does not affect the accuracy of SST variations along the core. In any case, this offset is probably small: *G. ruber* is a warm-water species and is especially abundant when the surface temperature is between 21 and 29 °C and surface salinity between 34.5 and 36‰ (ref. 25). As these conditions permanently prevail in the Mozambique Channel, we assume that *G. ruber* lives during the whole year. The temperature derived from isotope data for the Upper Holocene is 25.7 °C (Fig. 2b), which supports this hypothesis because the mean annual temperature of surface water is 26 °C²⁶.

Because the isotope method for estimating surface water temperature relies on the point-by-point subtraction of two independent records, errors are highly sensitive to small phase differences between the records and hence to the accuracy of the timescales. Taking account of an uncertainty of 0.1‰ owing to local changes in the surface seawater isotopes, although isotopic ratios are measured with a precision of 0.07‰ (equivalent to 0.35 °C), we do not expect an accuracy better than 1.5 °C from comparing the isotope record of core MD79-254 with that of mean seawater $\delta^{18}\text{O}$ record.

Over the past 135 kyr, isotope SST estimates (Fig. 2b) average 24.1 °C and vary by a maximum of 4 °C. SSTs similar to today (≥ 26 °C) occur only during the Holocene (at 10 kyr BP (before present)) and the lower part of isotope stage 5 (117 kyr BP). During the glacial phases, SSTs were 3 °C cooler than today. The temperature was never more than 4 °C colder than today. In the southwestern Indian Ocean, SSTs are high and relatively uniform in the warm season. Thus we interpret the changing SST as largely reflecting changes in the cold season (August).

Faunal sea surface temperature estimates

We (N.B. and W.L.P.) have also counted the relative frequency of planktonic foraminiferal species for the samples used in the isotope analysis. We identified 33 species ($>150\text{ }\mu\text{m}$) using the CLIMAP taxonomy² and counted an average of 355 individuals in each sample. In general, the fauna of MD79-254 is dominated by tropical and subtropical species, especially *G. ruber*. Using the factor assemblages defined by Hutson and Prell²⁷, the tropical assemblage accounted for almost 90% of the variance over the past 135 kyr. The periodic appearance of the *Globorotalia inflata* ($>4\%$ variance) reflects incursions of cooler, higher-latitude transitional waters into the Mozambique Channel.

The relative amounts of fauna were transformed into estimates of SST using the modern-analogue technique^{28–31}. We used a squared chord dissimilarity coefficient *D* to measure the similarity between samples. The range of *D* was 0–2, with the smaller values being most similar. The SST data for the five most similar

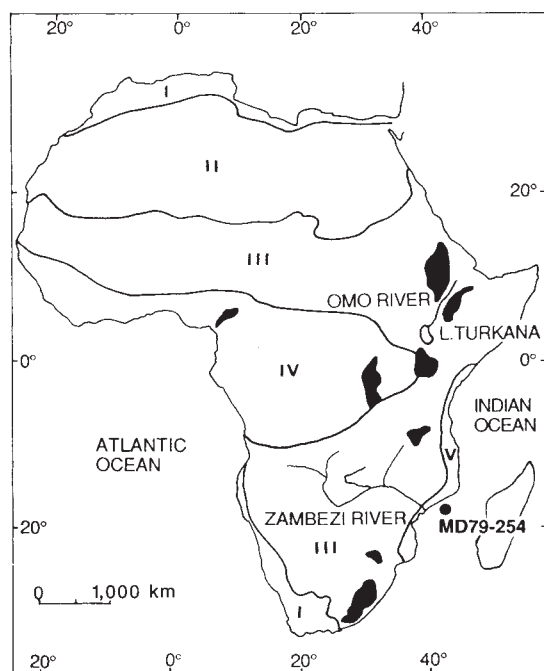


FIG. 1 Map of Africa showing location of core MD79-254, Omo river, main phytogeographical mapping units, after White⁴³ (I, Mediterranean; II, Saharo-Sindian; III, Sudano-Zambezian; IV, Guineo-Congolian; V, Zanzibar-Inhambane mosaic; black, Afromontane).

samples were averaged to obtain an SST estimate and standard deviation for each level. The Indian Ocean calibration data set has 290 samples for the analogue search. We found reasonable analogues for all levels with an average D of 0.13 (standard deviation, s.d., 0.02). The average warm-season (February) SST was 27.7 °C (s.d. = 0.95) and the average cold-season (August) SST was 24.4 °C (s.d. = 1.75). This relatively high standard deviation reflects the location of the core close to large SST gradients between subtropical and transitional waters and the possible occurrence of non-recent core tops in the global database, which could give slightly different temperatures for similar faunal assemblages. The Indian Ocean CLIMAP equation²⁷ gives almost identical results. These high standard deviations are common in areas with high SST gradients and may reflect the mismatch of modern SST data relative to the long-term average fauna. SSTs for the cold season (August) range from a high near 27 °C at ~10 kyr to a low of ~22.5 °C at ~25 and 48 kyr (Fig. 2c). Warm-season SSTs have a smaller range (28.5–26.5 °C). Overall, the warmest SST occurred during interglacial phases at ~10, 80 and 120 kyr and the coldest SST occurred during glacial stages 2 and 3 at ~18–50 kyr. Intermediate SSTs typify much of interglacial stage 5.

Pollen terrestrial surface temperature estimates

The pollen record of marine core MD79–254 has helped establish the fluctuations of vegetation and climate in southeast tropical Africa^{32,33}. The dominant pollen input is fluvial, as shown by the abundance of fern spores. We found pollen concentrations between 164 to 11,034 pollen grains per gram of sediment, with an average pollen count of 323 in each sample. A detailed study will be published elsewhere. In tropical Africa, temperature is the main control in the altitudinal zoning of vegetation. To estimate palaeotemperatures, we needed to

investigate the present-day transport pattern of key pollen types along the altitudinal profile of a river transect and the abundance of these pollen types relative to the surface temperature gradient. Such data are scarce in Africa, and there is no present-day pollen transect in a nearby river likely to be the present source for pollen found in core MD79–254. We used pollen from the Omo basin, situated in northeast tropical Africa as a calibration data set for the pollen in core MD79–254. The selected pollen types are identified botanically at the family level. We consider that the climate requirements for the plant taxa are similar enough north and south of the Equator to justify use of the Omo pollen data as a calibration data set for the pollen in our deep-sea core. We focused on *Combretaceae* and *Ericaceae* because their pollen is well represented in both the Omo samples and in the core, and because their general distribution patterns in East Africa show contrasting altitudinal and thermal affinities^{34,35}. *Combretaceae* are a family of trees, shrubs and vines found in a wide variety of habitats and generally associated with lowland forests and woodlands^{7,34}. In the highlands, the montane forest, ericaceous and alpine belts occur from lower to higher altitudes, *Ericaceae* being an important component of cold and dry high-altitude formations^{35,36}.

In Ethiopia, the Omo basin lies between latitudes 4°24' and 9°20' N and longitudes 35° and 38° E. The Omo river originates in the Ethiopian plateau west of Addis-Abeba and flows into Lake Turkana (Fig. 1). The mean annual temperature is 27 °C at the lake level and 15–20 °C on the plateau³⁷. The lowland vegetation type is sahelian with a dense riverine forest. Above 800 m, the vegetation grades to a sudanian type, well defined above 1,200 m. The valley is bordered by highlands exceeding 3,000 m. Montane forests occur between 1,500–2,000 m and 2,500–3,000 m, and ericaceous thicket and afro-alpine grassland occur above 3,000 m³⁸. In Mozambique, mean annual temperatures range from 25–26 °C on the coast to <20 °C in the highlands³⁹. The forest-bushland-grassland mosaic occupies a belt 50–200 km wide along the coast, and extends along broad river valleys. The coastal plain is backed by plateaus forming the largest part of the Zambezi domain. They lie above 900 m and are covered most extensively by woodland and savannah. The closest mountains are Mlanje (3,000 m) and Namuli (2,500 m) surrounding Lake Malawi. Most of the mountains lie in the forest belt, only Mt Mlanje is high enough to support a fragmentary ericaceous belt⁴⁰.

We (E.V.C.) used 29 pollen samples for the calibration. One sample is from Lake Wenchi (2,900 m) situated on the plateau near the Omo source, at the lower limit of the ericaceous belt⁴¹. Twenty-eight samples are muds from the river along a 890-km transect in the intermediate (1,100 m) and low (395 m) valley, from the Abalti bridge down to Lake Turkana (R. Bonnefille and G. Buchet, personal communication). Pollen transport and deposition is mostly dependent on the location and altitude of source areas. Potential source areas for *Ericaceae* and *Combretaceae* pollen are reflected in the distribution of their frequency along the transect. The temperature at each sampling site was calculated by interpolating the annual mean temperature data for western Ethiopia and Kenya⁴² at various altitudes. We used mean annual temperatures because the annual range of temperature, taken as the difference between the monthly maximum and minimum temperatures is very low (rarely exceeding 3 °C). Moreover, the coldest and warmest months are not always February and August, which would make any comparison with season SST difficult. The temperature range in the calibration data is 11.7 °C. Multiple linear regression of the percentage of *Combretaceae*, %C (range of 0.0–6.3 in calibration data) and the percentage of *Ericaceae*, %E (range of 0.0–16.0 in calibration data) on T yielded the following transfer function (correlation coefficient, $r = 0.94$; standard error = 0.70; number of data, $n = 29$; Fisher test = 113.59; probability of test < 10^{-4})

$$T = 24.69 + 0.39 \times \%C - 0.59 \times \%E \quad ^\circ\text{C}$$

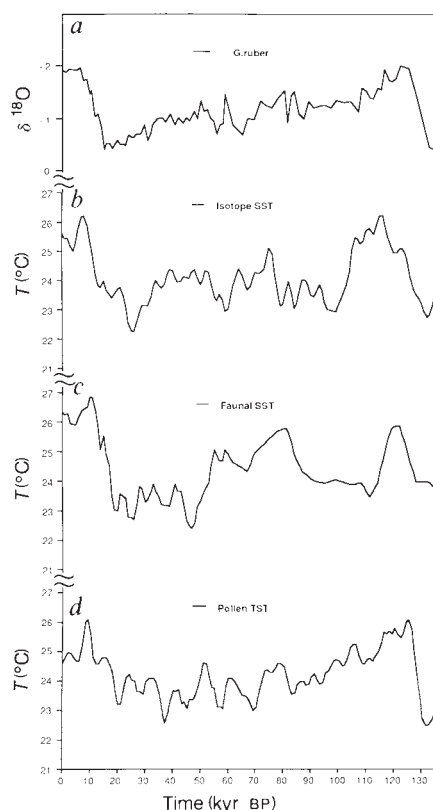


FIG. 2 a, Time series of oxygen isotope stratigraphy. b, Isotope SST estimates. c, Faunal SST estimates (cold). d, Pollen TST estimates. All the estimates are interpolated to 1-kyr intervals (smoothed by 3-point filters). $\delta^{18}\text{O} = [(^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{standard}}] - 1$, where the standard is PDB.

Application of this regression equation to the MD79–254 pollen data (range of %C and %E are 8.0 and 5.7, respectively) yields temperature estimates that apply to the closest terrestrial area, which is the Mozambique coast. All but one of the fossil samples fit the ranges of %C and %E in the calibration data. Although we need more samples from other rivers, and need to test our transfer function by applying it to modern river sediments outside the part of northeast Africa used for calibration, this transfer function provides useful preliminary estimates which indicate a TST (Fig. 2) slightly above 26 °C during the Holocene (9 kyr) and the last interglacial peak (125 kyr). During the past glacial period, TSTs ranged from 22 to 23 °C with significant variability over much of the record. The average TST over the past 135 kyr was 24.1 °C and the glacial-to-interglacial TST range for the Mozambique coast exceeds 3.5 °C.

Discussion and conclusion

Before comparing these palaeotemperature estimates, we stress that they are independent estimates, each with quite different assumptions, data and sensitivities. We assume that each system (isotopes, plankton and pollen) responds to the regional climate but with different sensitivities and different recording mechanisms. Hence, we expect correlation between estimates only for large climate changes and not necessarily for smaller-scale changes, which certainly must reflect the different environmental settings and processes. Given this caveat, we find that the estimates of mean temperature over the past 135 kyr fall within 0.5 °C of each other (24.1–24.3 °C) and that the maximum temperature changes of ~3.5 to 4 °C are similar for each record. Thus, the terrestrial and marine estimates of surface temperature are almost identical in their average value and range of temperature change.

More detailed comparison of the three temperature records reveals both similarities and differences. The most striking similarities are in the temperature changes associated with the glacial–interglacial transitions. Although each record has its own distinctive pattern, all three records indicate lower temperatures during glacial phases and a rapid increase associated with deglaciation, and higher temperatures during the Holocene and last interglacial peak (~120 kyr). There are, however, several interesting differences. During the previous deglaciation (termination II), all records show a rapid temperature increase to a maximum

near the peak of isotope stage 5.5, but the duration of high temperatures is distinctly longer in isotope-based estimates than in faunal⁴ and pollen-based estimates. The records are less coupled in the interglacial periods. Both marine records indicate cooler temperatures during the middle of isotope stage 5 and warmer temperatures in the upper stage 5. But the timing of the two records is clearly different. Likewise, the TST record has some common structure with the marine records but with variable timing. In some instances, the records are completely out of phase.

The detailed interpretation of these records must take into account that a number of mechanisms are acting to give the actual terrestrial and marine temperatures and that other environmental processes are involved in preserving these proxy temperature data. In addition to direct radiation budgets, the marine estimates reflect the strength of the Agulhas current (stronger transport of tropical elements to the site), the latitude of the high winter SST gradient to the south of the site, and possible changes in the evaporation–salinity relationship. TSTs reflect the precipitation regime (stronger or weaker monsoon influence) as well as the general atmospheric circulation and fluvial transport. The timing of small-scale temperature changes is then complicated by each recording system and detailed comparisons between land and sea records may be difficult even in ideal cases.

The discussion about the inconsistencies between terrestrial and marine palaeotemperature reconstructions for the Last Glacial Maximum is not closed. The faunal evidence in core MD79–254 indicates that the Last Glacial Maximum was 3 °C colder than today, whereas the CLIMAP data from off the shore of East Africa^{1,2} indicate a difference of only 1 °C. But, as already noted, coastal SST may be subject to local and boundary effects not recorded farther off shore. Despite the different processes involved, both lowland terrestrial and marine estimates of surface temperature at this site seem accurately to record the average value and the range of temperature changes over major glacial–interglacial cycles. Our study does not reveal a discrepancy between land and sea temperature. This method can be applied to other sites. Our comparison of terrestrial and marine temperature estimates is the most detailed available so far, and the results indicate that the relationship between terrestrial and marine temperatures may be solved by this kind of approach. □

Received 21 May; accepted 14 September 1990.

- CLIMAP Science **191**, 1131–1137 (1976).
- CLIMAP project members *Geol. Soc. Am. Map Chart Ser.* MC-36 (1981).
- Peterson, G. M. *et al. Quat. Res.* **12**, 47–82 (1979).
- Webster, P. & Streten, N. *Quat. Res.* **10**, 279–309 (1978).
- Hastenrath, S. *The Glaciers of Equatorial East Africa* (Reidel, Dordrecht, 1984).
- Rind, D. & Peteet, D. *Quat. Res.* **24**, 1–22 (1985).
- Hamilton, A. C. *Environmental History of East Africa*, 285 (Academic, London, 1982).
- Van Zinderen Bakker, E. M. & Coetzee, J. A. *Rev. Palaeobot. Palynol.* **55**, 155–174 (1988).
- La Fontaine, C. V. thesis, Univ. Wisconsin (1988).
- Kutzbach, J. E. & Guetter, P. J. *J. Atmos. Sci.* **43**, 1726–1759 (1986).
- Bonnefille, R. *et al. Nature* **346**, 347–349 (1990).
- Duplessy, J. C. in *Climatic Change* (ed. Gribbin, J.) 46–67 (Cambridge University Press, 1978).
- Emiliani, C. *J. Geol.* **63**, 538–578 (1955).
- Prell, W. L. *et al. Palaeogeography* **1**, 137–162 (1986).
- Imbrie, J. *et al. in Milankovitch and Climate* (eds Berger, A. & Imbrie, J.) 269–305 (Reidel, Dordrecht, 1984).
- Emiliani, C. *Science* **173**, 1122–1124 (1971).
- Shackleton, N. J. & Vincent, E. *Marine Micropaleont.* **3**, 1–13 (1978).
- Duplessy, J. C. *C. r. hebdom. Seanc. Acad. Sci.* **271**, 1075–1078 (1970).
- Cullen, J. L. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 315–356 (1981).
- Labeyrie, L. *et al. Nature* **327**, 477–482 (1987).
- Shackleton, N. J. in *Colloque Int. CNRS no 219, Variation du Climat au Cours du Pleistocène* (ed. Labeyrie, J.) 203–209 (CNRS, Paris, 1974).
- Duplessy, J. C. *et al. in Long and Short Term Variability of Climate* (eds Wanner, H. & Siegenthaler, U.) 83–116 (Springer, Berlin, 1988).
- Duplessy, J. C. *et al. Palaeogeogr. Palaeoclimatol. Palaeoecol.* **33**, 9–46 (1981).
- Duplessy, J. C. *et al. Science* **213**, 1247–1250 (1981).
- Bé, A. W. H. & Hutson, W. H. *Micropaleontology* **23**, 369–414 (1977).
- Wyrtki, K. *Oceanographic Atlas of the International Indian Ocean Expedition* (NSF, Washington, DC, 1971).
- Hutson, W. H. & Prell, W. L. *J. Paleont.* **54**, 381–399 (1980).
- Hutson, W. H. *Science* **207**, 64–66 (1979).
- Prentice, I. C. *Rev. Palaeobot. Palynol.* **31**, 71–104 (1980).
- Overpeck, J. T. *et al. Quat. Res.* **23**, 87–108 (1985).
- Prell, W. L. *The stability of Low-Latitude Sea-Surface Temperatures*, Rep. TR025 (U.S. Department of Energy, Washington DC, 1985).
- Livingstone, D. A. *Limnol. Oceanogr.* **16**, 349–356 (1971).
- Meadow, M. E. *Palaeoecol. Afr.* **16**, 353–390 (1984).
- Wickens, G. E. in *Flora of East Africa* (ed. Polhill, R. M.) 99 (Ministry of Overseas Development, London, 1973).
- Hedberg, O. *Svensk bot. Tidskr.* **45**, 141–202 (1951).
- Hedberg, O. *Symb. bot. Upsal.* **15**, 5–412 (1957).
- Chaffey, D. R. *South-West Ethiopia Forest Inventory Project Rep.* 31, 315 (Ministry of Overseas Development, London, 1979).
- Pichi-Sermolli, R. E. G. *Webbia* **13**, 15–132 (1957).
- Leroux, M. *Le Climat de l'Afrique Tropicale* (CNRS, Paris, 1983).
- White, F. in *Biogeography and Ecology of Southern Africa* (ed. Werger, M. J. A.) 463–513 (Junk, The Hague, 1978).
- Bonnefille, R. & Buchet, G. *Mém. Trav. Ecole Pratique des Hautes Etudes* **17**, 143–158 (1987).
- Ruddolf, W. *World Climates* (Wissenschaftliche, Stuttgart, 1981).
- White, F. *The Vegetation of Africa* (UNESCO, Paris, 1983).

ACKNOWLEDGEMENTS. We thank T. Atkinson, R. Bonnefille, G. Buchet, J. Maley and J. T. Overpeck for comments, D. Duzer for technical and pollen analysis assistance, and B. Le Coat and J. Antignac for help with the isotope analyses. This work was supported in France by the Centre National de la Recherche Scientifique, the Commissariat à l'Energie Atomique and the Programme National d'Etude de la Dynamique du Climat, and in the United States by the NSF.

Isolation of an endogenously processed immunodominant viral peptide from the class I H-2K^b molecule

Grada M. Van Bleek & Stanley G. Nathenson

Department of Cell Biology and the Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

Using an approach for isolating and characterizing peptide fractions that are intracellularly associated with major histocompatibility complex class I molecules, the major peptide recognized by cytotoxic T cells specific for the vesicular stomatitis virus has been isolated from the H-2K^b molecule of infected cells. This endogenously processed octapeptide is allele-specific as it does not bind to H-2D^b molecules, and contains the core sequence of the epitope of the nucleocapsid protein of the vesicular stomatitis virus identified by testing with exogenous synthetic peptides.

CYTOTOXIC T lymphocytes (CTL) recognize foreign antigens in association with class I major histocompatibility complex (MHC) molecules¹. The current view of antigen presentation by class I molecules is that endogenously synthesized viral or other protein antigens are degraded in the cytoplasm into small peptides which translocate into a pre-Golgi compartment, and interact with class I heavy chains, triggering proper folding and association with β_2 -microglobulin (β_2m)²⁻¹⁵. The complex is then transported to the cell surface^{4,8}. Considerable amounts of data on allele-specific motifs of antigenic peptides have been derived using CTL recognition assays of targets sensitized by exogenously added synthetic peptides¹⁶⁻²⁶. However, features of such peptides might differ considerably from endogenously processed peptides^{27,28}. Therefore we developed a method for the isolation and characterization of endogenously synthesized peptides which allows the analysis of the number, size and types of peptides that have been processed and have become physically associated with MHC molecules.

Class I molecules loaded with vesicular stomatitis virus (VSV) peptides were isolated from infected cells by immunoprecipitation at a time during infection when the synthesis of most proteins other than viral and MHC products is shut down. The immunoprecipitates were denatured and fractionated according to size, and the fraction of low relative molecular mass (M_r) was analysed by HPLC, which revealed a unique K^b-associated peptide peak for infected cells but not for uninfected cells. This peptide was established as the major VSV antigenic determinant by both its sequence and biological activity.

The major K^b-restricted N peptide

The CTL response against VSV is exclusively K^b-restricted in H-2^b mice (K^b, D^b) and is directed against the nucleocapsid protein (N)²⁹⁻³⁰, presumably against a peptide in the N-terminal one-third of the molecule (K. Ishikawa and D. S. Lyles, unpublished results).

To identify this epitope, a set of 75 overlapping peptides (13 amino acids in length) was synthesized³¹, overlapping by 11 amino acids and spanning the N-terminal 161 residues of the protein. Four overlapping peptides spanning residues 47 to 65 sensitized target cells to recognition by CTL clone 33 specific

for VSV (Fig. 1). The minimum requirement for biological activity in the sensitization assay was thus determined to be the core sequence: GYVYQGL (single-letter amino-acid code; N53-59). Because a clone was used in these experiments and the set of peptides tested represented only the N-terminal one-third of the molecule, it was possible that other peptides from N also could contribute to the recognition by bulk CTL. To exclude this possibility, EL4 cells treated with N53-59 peptide were used as cold targets to inhibit lysis of N1 cells (EL4 cells, transfected with the complete VSV N gene). Cells coated with a peptide containing N53-59 completely inhibited lysis of N1 cells (Fig. 2). The N53-59 peptide is therefore the dominant antigenic epitope in VSV-specific K^b-restricted recognition.

Major peptide associated with K^b is viral

EL4 cells infected with 30 plaque-forming units per cell of VSV Indiana strain (VSV-Ind.) begin synthesizing viral protein 2.5 h after infection and begin to considerably decrease the synthesis of cellular protein 5 h after infection (post infection, p.i.); synthesis of K^b and D^b continues at relatively normal levels until

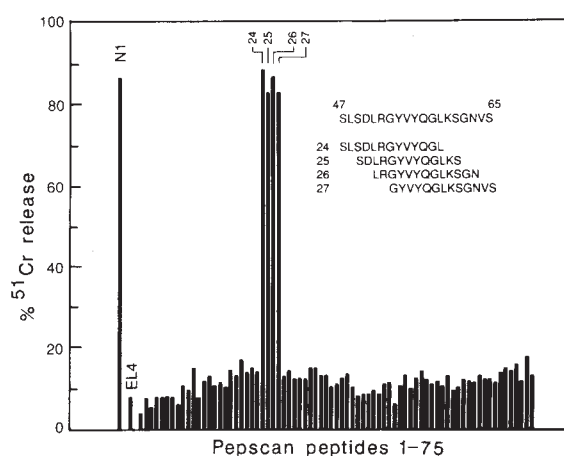


FIG. 1 CTL clone 33 recognizes four overlapping peptides (N47-59, N49-61, N51-63, N53-65), sharing a minimal core of seven amino acids (N53-59, GYVYQGL) out of a panel of 75 overlapping peptides spanning the N-terminal 161 amino acids of the nucleocapsid protein of VSV.

METHODS. Peptides were synthesized according to the protein sequence of N protein from VSV-Ind as reported by Gallione *et al.*³², by an automated pepsan method as described^{41,42}, and detached from the solid phase by acid treatment³¹. Clone 33 was maintained by weekly restimulations with irradiated B6-H-2^{b/m8} (bm8) spleen cells as described previously³⁰. ⁵¹Cr release (Na₂⁵¹CrO₄) assays were essentially as described by DeWaal *et al.*⁴³. Peptides were added dissolved in PBS (50 μ l) to 2,500 EL4 cells (100 μ l assay medium: RPMI+10% FCS) in 96-well plates and were present throughout the 4 h. ⁵¹Cr release: ⁵¹Cr-labelled EL4 and N1 cells were exposed to 1.25 $\times 10^5$ clone 33 cells (effector to target (E/T) ratio, 5:1; 50 μ l) in a final volume of 200 μ l. N1 is an EL4 cell transfected with the nucleocapsid protein gene of VSV (ref. 29). ⁵¹Cr release was measured, after 4 h of contact, in 150 μ l supernatants. Percentage specific lysis was calculated as: 100 $\times$ (release by CTL - release in medium alone) / (2.5% Triton X-100 release - release in medium alone).

7 h p.i. (data not shown). VSV-infected EL4 cells were biosynthetically radiolabelled (2.5–7 h p.i.) using ^3H -labelled amino acids present in the N53–59 peptide. K^b and D^b were isolated by immunoprecipitation from detergent-lysed cells, and the low- M_r fraction containing peptides of 7–15 amino acids was retrieved from the class I/peptide complex by acid denaturation followed by Centricon-10 centrifugation. The peptide fraction was further separated by reverse-phase HPLC (Fig. 3). Similar peptide profiles were obtained from the K^b immune precipitates derived from uninfected and infected EL4 cells (plotted in Fig. 3a as separate runs in the same graph). But at 23.2 min after injection, a prominent extra peak was clearly present in the (^3H)Arg-, (^3H)Leu-labelled) peptide fraction eluted from K^b isolated from VSV-infected cells. Preparations from cells doubly labelled with (^3H)Tyr and (^3H)Lys, and with (^3H)Tyr and (^3H)Leu gave similar results (data not shown). In addition to the virus peak, two smaller peaks not present in uninfected EL4 cells were observed in the K^b profile at 31.2 min and 33.5 min respectively. These peptides did not have the sequence pattern of the 23.2-min peak (data not shown). The D^b -associated peptides of VSV-infected and uninfected EL4 cells (Fig. 3b) were clearly different from the K^b -associated peptides and, furthermore, the virus peptide peak present in the K^b profile was absent in the D^b profile. However, the D^b profile of infected EL4 cells also contained a major extra peak eluting at 31.2 min that was not present in uninfected cells and was different from the K^b -associated virus peak. The D^b -associated virus peak seemed to coelute with a minor virus peak in the K^b profile, but preliminary sequencing showed that they were different (not shown).

The HPLC described above was performed on the fraction of peptides of M_r 1,000–1,500 (1–1.5K) recovered from K^b . A rough estimate of the relative ratio of peptides to heavy chain in this fraction was about 0.5 to 1 as deduced by the relative amounts of radioactivity in these fractions when a complete (^3H)Tyr-, (^3H)Leu-labelled immunoprecipitate was separated under denaturing conditions on Sephacryl S-200 HR (not

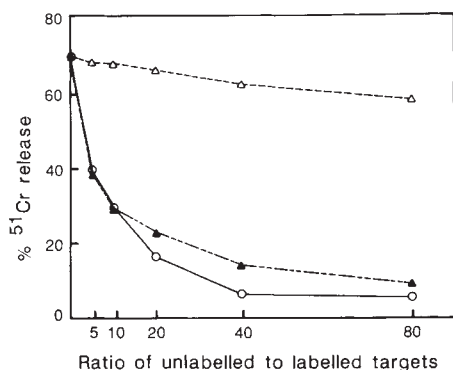


FIG. 2 EL4 cells treated with peptide N53–59 can completely block a bulk CTL response against N1 cells. E/T ratio, 50:1. Cold targets: N1, $\circ$ – $\circ$; EL4, $\triangle$ – $\triangle$; EL4 plus peptide N53–59, $\blacktriangle$ – $\blacktriangle$.

METHODS. Effector cells were prepared by co-culturing 30×10^6 spleen cells of C57BL/6 (B6, H-2^b) mice that had been immunized 3 weeks earlier intraperitoneally with 10^6 PFU of VSV, with 3×10^6 irradiated (20,000 rad) N1 cells for 5 days in 10 ml RPMI plus 10% FCS, 100 IU/ml⁻¹ penicillin, 100 μg ml⁻¹ streptomycin, 2×10^{-5} M 2-mercaptoethanol, 2 mM L-glutamine in upright 25 ml flasks at 37 °C in 5% CO₂. Various numbers of EL4 cells pretreated for 30 min at 37 °C with 500 nM peptide N53–59 were added in 50 μl (in 500 nM peptide solution) to 5×10^4 effector cells in 50 μl RPMI plus 10% FCS in 96-well microtitre plates. Then 2×10^3 ^{51}Cr -labelled N1 cells were added in 100 μl medium and cells were incubated for 4 h at 37 °C in humidified air and 5% CO₂. Thereafter ^{51}Cr release was measured in 150 μl of supernatants. Similarly, untreated EL4, and N1 cells were used as negative and positive controls, respectively. Percentage specific lysis was calculated as described in the legend to Fig. 1. Spontaneous ^{51}Cr release of N1 cells was 7.4%. Similar results were obtained in three independent experiments.

shown). Because the viral peptide derived from K^b amounts to about 10% of the 1–1.5K fraction as deduced from the HPLC profile, at least 5% of the isolated radioactive K^b molecules are occupied by viral peptides. It is of interest that in both D^b and K^b profiles, aside from virus specific peaks, there are about 8 to 10 prominent peaks which seem to represent self peptides, defined as MHC-associated peptides present in both the virus-infected and the uninfected cells.

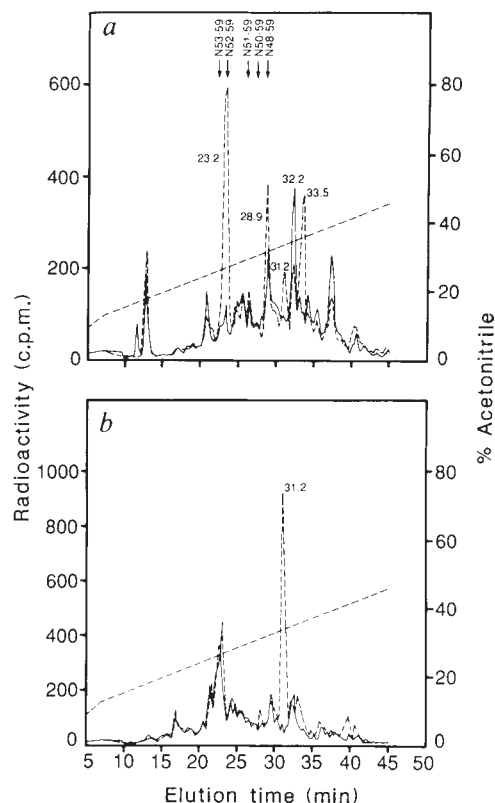


FIG. 3 HPLC analysis of the radiolabelled peptide fractions of MHC class I K^b (a) and D^b (b) molecules. ($\cdots$), Peptides from infected EL4 cells; (—), peptides from uninfected cells; (---) acetonitrile gradient.

METHODS. EL4 cells (13×10^6) were incubated with 30 PFU per cell VSV-Ind for 1 h at 37 °C in 35 ml DMEM medium (Gibco). Cells were then washed once, resuspended in 60 ml DMEM, 10% FCS and further incubated for 1.5 h at 37 °C. Two-and-a-half hours after infection cells were washed once in minimal essential medium (MEM, Gibco) without Arg and Leu (MEM–R,L), then incubated for 4.5 h on a roller wheel: 10^8 cells in 12 ml MEM–R,L supplemented with 5% dialysed FCS, 0.5 mCi [^3H]Arg and 1 mCi [^3H]Leu. Similarly, 13×10^6 uninfected EL4 cells were labelled. After labelling, cells were washed once in PBS and lysed in 10 ml of 20 mM Tris–HCl, pH 7.5; 150 mM NaCl, 0.5% Nonidet P-40 and a proteolytic enzyme inhibitor mix containing 5 μg ml⁻¹ Aprotinin, 10 μg ml⁻¹ Leupeptin, 10 μg ml⁻¹ Pepstatin (all from Boehringer), 100 μM iodoacetamide, 3 ng ml⁻¹ EDTA, 0.2% azide and 2 mM PMSF. Immunoprecipitation of Nonidet P-40 (NP40)-solubilized cell lysates was done using in order of sequence, normal mouse serum, D^b -specific antibody B22.249 and K^b -specific antibodies; B8.24.3 (ref. 44) and Y3 (ref. 45). After seven washes, the immunoprecipitate was denatured by boiling for 3 min in 10% acetic acid and was separated into a high- M_r , heavy chain and β 2m fraction and a peptide fraction by Centricon 10 (Amicon, Danvers, Maryland) centrifugation. The low- M_r material, when eluted on a Sephadex-g50.40 column in 10% acetic acid, appears to contain 90% of the available material of M_r 1–1.5K in the immunoprecipitate as determined by a parallel run on unfractionated material (not shown). This material was analysed by HPLC on a mini reverse-phase column (RP.8 100 $\times$, 2.1 mm, 5 μm ; Brownlee) by using Hewlett Packard (HP1090) equipment. The gradient was made from solution A (H_2O with 0.1% trifluoroacetic acid), and solution B (acetonitrile with 0.1% trifluoroacetic acid). Flowrate, 0.25 ml min⁻¹; two-drop fractions collected. Radioactivity was measured in an aliquot of one-fifth of each fraction. The markers are noted at the top of a and are a nested set of synthetic peptides in the N-protein sequence: $^{49}\text{SDLRGVYVQGL}^{59}$, all with Leu 59 as C terminus and shortened from the N terminus to the C terminus, ending with peptide N53–59 of seven amino acids.

Activity of the K^b-bound VSV octamer

The peptide in the peak eluting at 23.2 min was biologically active because it sensitized EL4 cells to lysis by VSV-specific K^b-restricted bulk CTL (Fig. 4a). A nonvirus-specific (self) peak (28.9 min) and the D^b virus-specific peak (31.2 min) were inactive in the same test. The 23.2-min peak was also active in sensitizing EL4 cells when tested with clone 33 (Fig. 4b).

To identify the peak eluting at 23.2 min, material from two different metabolic labelling experiments was sequenced. The peptide containing both [³H]Tyr and [³H]Leu showed a strong signal at positions 3 and 5 and a lower signal at cycle 8 (Fig. 5a). In a second experiment in which ³H Tyr and ³H Lys labels were used, only positions 3 and 5 were positive, identifying these two residues as Tyr (N54, N56; (Fig. 5b). Because in the second experiment there was no signal at position 8, Leu must occupy that position (N59). The positions of radiolabelled Tyr and Leu residues agree with the amino-acid sequence found in the peptide-scanning experiment and positions the N terminus at Arg 52. From Fig. 5b it is seen that a [³H]Lys signal at position 9 is absent, suggesting that the natural epitope from VSV-N protein starts at Arg 52 and extends to Leu 59. This was confirmed in an experiment showing that the synthetic octamer N52-59 elutes at 23.2 min in the HPLC profile—the same position as the virus peak retrieved from K^b—whereas synthetic peptides N53-59, N51-59, N50-59, N48-59 all eluted at different times. Therefore, the natural epitope for the K^b-restricted VSV-specific CTL response is the N protein-derived octamer, N52-59.

Discussion

Because exogenously added sensitizing synthetic peptides may differ from their physiologically relevant peptides^{27,28} and it is unclear how relevant the various binding motifs¹⁶⁻²⁶ may be, we developed a method to isolate peptides that are made by intracellular processing pathways and are endogenously bound to MHC class I molecules. This approach enabled us to identify the naturally produced peptide derived from the N protein of VSV that binds to the H-2K^b class I molecule. The use of radiolabelled amino acids produced high sensitivity and assured that the peptides retrieved were actually biosynthetic products. These experiments allowed us to demonstrate that the endogenously processed peptide derived from the VSV N protein contained the same core sequence as the synthetic viral epitope established using an assay in which peptides were added exogenously.

The viral peptide that we isolated from the K^b molecule was the most prominently labelled peptide associated with K^b in the VSV-infected EL4 cell. But in addition to the major virus peak, two smaller peaks (31.2 min, 33.5 min) that were absent in unin-

fected cells may be virus-derived. However, they do not contribute significantly to the bulk VSV CTL response as cold target blocking experiments showed that EL4 cells treated with peptide N53-59 could completely block a bulk CTL response of VSV-N-specific CTL against N1 target cells which contain the complete VSV-N gene. Thus the most abundant peptide of the HPLC profile, representing about 5 to 10% of the radioactivity in the peptide fraction, was also the major epitope for CTL. These results are consistent with earlier reports showing that from a whole virus only one dominant peptide is recognized per antigen-presenting molecule^{11,12,33-35}. The major antigenic sequence was recovered in a discrete peak in the HPLC profile representing an octamer. This finding contrasts with a study describing the isolation of processed peptides of hen egg lysozyme-pulsed cells in which a heterogeneous set of overlapping hen egg lysozyme peptides of different lengths was found to be associated with class II molecules³⁶. The difference between the results might reflect actual differences in proteolytic activity in class I and class II processing pathways; that is, class II-associated peptides may result from enzymatic cleavage with many different proteases or with proteases with broad substrate specificity, whereas class I processing may use a more select set of proteases.

Our findings also contrast with those from studies which used synthetic peptides to block specific recognition by CTL which showed that although there was allele specificity, 20-40% of arbitrarily chosen sets of peptides could actually bind to MHC allelic products^{16,38}. This high level of arbitrary binding was not found in our experiments. From the five VSV proteins available for processing and presentation (N, NS, M, G and L) only four processed peptides bound, three of them to K^b and one to D^b.

That the dominant K^b-restricted antigenic peptide was bound only to the K^b and not the D^b molecule suggests a high degree of specificity in the binding of peptides to specific MHC allelic products; this contrasts with the findings of *in vitro* approaches using an assay of binding MHC molecules to peptide-coated plates, which have suggested a promiscuous binding of many peptides to all class I molecules³⁷. Although it was not biologically active in the context that we examined, the D^b-associated viral peptide also showed highly specific binding.

The profiles of self peptides derived from K^b and D^b were dissimilar and the number of major peptides represented by these profiles seemed surprisingly few considering the enormous reservoir and the continuous turnover of numerous cellular proteins in the cytoplasm. Because endogenous peptide repertoires seem more select than could be expected from binding studies with synthetic peptides, other factors in addition to mere structural requirements for binding must influence the repertoire of peptides that can be functionally presented by class I MHC molecules. For example, processing has already been reported

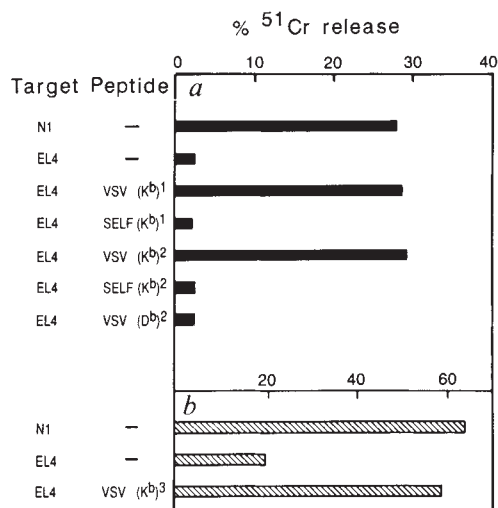


FIG. 4 Material eluting at 23.2 min in the HPLC profile of peptides isolated from K^b of VSV-infected EL4 cells can sensitize EL4 cells for recognition by N-specific CTL. 1, Peptides from a [³H]Tyr, [³H]Leu labelling experiment; 2, peptides from [³H]Arg, [³H]Leu labelling; 3, peptide from [³H]Tyr, [³H]Lys labelling.

METHODS. Peak fractions in the K^b and D^b profile of VSV infected EL4 cells eluting at 23.2 min (26% acetonitrile) were collected and lyophilized. The peptides were resolubilized in 250 µl PBS. As negative controls a presumably cellular (self) peptide eluting at 28.9 min (31% acetonitrile) in the [³H]Arg, [³H]Leu K^b profile (Fig. 3b) of VSV-infected cells and the viral peptide in the D^b profile eluting at 31.2 min (33% acetonitrile) as well as a prominent cellular peptide from the [³H]Tyr, [³H]Leu labelling experiment were used. These peptide solutions (50 µl) were added to 2,000 ⁵¹Cr-labelled EL4 cells in 100 µl RPMI plus 10% FCS. Then effector cells of bulk anti-N CTL prepared as described in the legend to Fig. 2 at E/T ratio 50:1 (a) or clone 33 (E/T ratio 5:1, b) were added, and after 4 h ⁵¹Cr release was measured in supernatants. Spontaneous ⁵¹Cr release was for all targets in both experiments <14%.

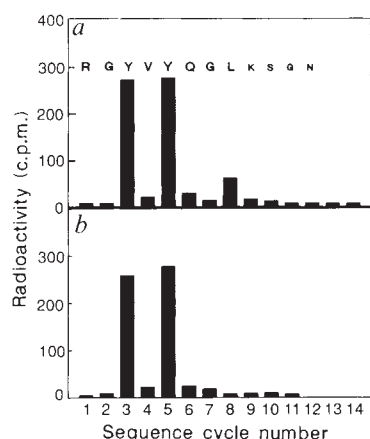


FIG. 5 Partial radiolabel sequence analysis of the peptide peak derived from K^b of VSV-infected EL4 cells eluting at 23.2 min after injection on HPLC. *a*, Metabolic labels [³H]Tyr, [³H]Leu; *b*, metabolic labels [³H]Tyr, [³H]Lys.

METHODS. Metabolic labelling and immunoprecipitations were as described in the legend to Fig. 3. Samples of the low-*M_r* fraction after Centricon centrifugation were concentrated in 30 µl. Peptides were sequenced on an Applied Biosystems 477A pulsed liquid sequencer, with fraction collection of the phenylthiohydantoin amino acids. Radioactivity was measured in the amino-acid fractions. The sequence of the synthetic epitope of N protein was aligned with the positions of the radioactive signals. Large letters, residues contained in the natural epitope; smaller letters, additional amino acids present in the N-protein sequence that are not included in the natural epitope.

as being able to influence the availability of potentially functional antigenic peptides^{27,28,39}. It is of interest that peptides of 7 to 13 amino acids, containing the antigenic core sequence, could all sensitize targets for CTL recognition (data not shown). But only the octamer which has positively charged amino acids on the C-terminal sides of the presumed cleavage sites was recovered *in vivo*. This may be relevant to the selectivity of the protein cleavage process and to the type of protease(s) involved in processing.

Certain peptides cannot be functionally recognized by CTL even though they do bind to MHC class I molecules, because they mimic self peptides to which tolerance may have been induced⁴⁰. An example for this phenomenon could be the absence of a D^b-restricted component in an anti-VSV CTL response in H-2^b mice, even though D^b seems to be able to associate with a peptide that was exclusively and abundantly present in VSV-infected EL4 cells and not in uninfected cells (Fig. 3).

Class I binding studies previously described use either enzymatic digests of proteins or synthetic peptides exogenously added to isolated MHC molecules or MHC molecules on intact cells and subsequently tested for biological activity of the peptide-MHC complex. These methods have proved useful in searching for non-self and potentially for self peptides presented by class I molecules. A recent study has described the characterization of physiologically produced peptides isolated from cell extracts which can be recognized by CTL, a procedure that has the potential to identify naturally produced antigenic peptides⁴⁶. The approach described here isolates intracellularly processed peptides directly from MHC class I molecules and provides the opportunity to obtain information about the lengths of natural peptides and the relative abundance of peptides binding to class I molecules, and further permits screening for differences between peptide profiles bound to class I molecules of different cell types or different allelic products. □

Received 3 August; accepted 24 September 1990.

1. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51–177 (1979).
2. Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457–467 (1985).
3. Townsend, A. R. M. *et al. Nature* **340**, 443–448 (1989).
4. Cox, J. H., Yewdell, J. W., Eisenlohr, L. C., Johnson, P. R. & Bennink, J. R. *Science* **247**, 715–718 (1990).
5. Yewdell, J. W. & Bennink, J. R. *Science* **244**, 1072–1075 (1989).
6. Nuchtern, J. G., Bonifacio, J., Biddison, W. E. & Klausner, R. *Nature* **339**, 223–226 (1989).
7. Lie, W.-R. *et al. Nature* **344**, 439–441 (1990).
8. Hosken, N. A. & Bevan, M. J. *Science* **248**, 367–370 (1990).
9. Cerundolo, V. *et al. Nature* **345**, 499–452 (1990).
10. Bastin, J., Rothbard, J. B., Davey, J., Jones, I. & Townsend, A. R. M. *J. exp. Med.* **165**, 1508–1523 (1987).
11. Braciale, T. J. *et al. J. exp. Med.* **166**, 678–692 (1987).
12. Oldstone, M. B. A., Whitton, J. L., Lewicki, H. & Tishon, A. *J. exp. Med.* **168**, 559–570 (1988).
13. Bjorkman, P. J. *et al. Nature* **329**, 506–512 (1987).
14. Bjorkman, P. J. *et al. Nature* **329**, 512–518 (1987).
15. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692–696 (1989).
16. Shimojo, N., Maloy, W. L., Anderson, R. W., Biddison, W. E. & Coligan, J. E. *J. Immun.* **143**, 2939–2947 (1989).
17. Takahashi, H. *et al. J. exp. Med.* **170**, 2023–2035 (1989).
18. Maryanski, J. L., Verdini, A. S., Weber, P. C., Salemme, F. R. & Corradin, G. *Cell* **60**, 63–72 (1990).
19. DeLisi, C. & Berzofsky, J. A. *Proc. Natn. Acad. Sci. U.S.A.* **82**, 7048–7052 (1985).
20. Rothbard, J. B. & Taylor, W. R. *EMBO J.* **7**, 93–100 (1988).
21. Guillet, J.-G., Lai, M.-Z., Briner, T. J., Smith, J. A. & Geftter, M. L. *Nature* **324**, 260–262 (1986).
22. Sette, A. *et al. Nature* **328**, 395–399 (1987).
23. Ogasawara, K., Maloy, W. L., Beverly, B. & Schwartz, R. H. *J. Immun.* **142**, 1448–1456 (1989).
24. Bhayani, H. & Paterson, Y. *J. exp. Med.* **170**, 1609–1625 (1989).
25. Sette, A., Buus, S., Colon, S. M., Miles, C. & Grey, H. M. *J. Immun.* **141**, 45–54 (1988).
26. Sette, A. *et al. J. Immun.* **143**, 3289–3294 (1989).

27. Fox, B. S., Carbone, F. R., Germain, R. N., Paterson, Y. & Schwartz, R. H. *Nature* **331**, 538–540 (1988).
28. Bodmer, H. C., Gotch, F. M. & McMichael, A. J. *Nature* **337**, 653–655 (1989).
29. Puddington, L., Bevan, M. J., Rose, J. K. & Lefrançois, L. *J. Virol.* **60**, 708–717 (1986).
30. Sheil, J. M., Bevan, M. J. & Lefrançois, L. *J. Immun.* **138**, 3654–3660 (1987).
31. Van der Zee, R., Van Eden, W., Melen, R. H., Noordzij, A. & Van Embden, J. D. A. *Eur. J. Immun.* **19**, 43–47 (1989).
32. Gallione, C. J., Greene, J. R., Iverson, L. E. & Rose, J. K. *J. Virol.* **39**, 529–535 (1981).
33. Taylor, P. M., Davey, J., Howland, K., Rothbard, J. B. & Askonas, B. A. *Immunogenetics* **26**, 267–272 (1987).
34. Bodmer, H. C., Pemberton, R. M., Rothbard, J. B. & Askonas, B. A. *Cell* **52**, 253–258 (1988).
35. Kast, W. M. *et al. Cell* **59**, 603–614 (1989).
36. Demotz, S., Grey, H. M., Appella, E. & Sette, A. *Nature* **342**, 682–684 (1989).
37. Bouillot, M. *et al. Nature* **339**, 473–475 (1989).
38. Carreno, B. M., Anderson, R. W., Coligan, J. E. & Biddison, W. E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3420–3424 (1990).
39. Schild, H., Rotzschke, O., Kalbacher, H. & Rammensee, H.-G. *Science* **247**, 1587–1589 (1990).
40. Nikolic-Zugic, J. & Bevan, M. J. *Nature* **344**, 65–67 (1990).
41. Geysen, H. M., Melen, R. H. & Barteling, S. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3998–4002 (1984).
42. Geysen, H. M., Barteling, S. J. & Melen, R. H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 178–182 (1985).
43. DeWaal, L. P., Kast, W. M., Melvold, R. W. & Melief, C. J. M. *J. Immun.* **130**, 1090–1096 (1983).
44. Kohler, G., Fischer-Lindahl, K. & Heusser, C. in *The Immune System*, 202–208 (Karger, Basel, 1981).
45. Jones, B. & Janeway, C. A. *Nature* **292**, 547–549 (1981).
46. Rotzschke, O., Falk, K., Wallny, H.-J., Faath, S. & Rammensee, H.-G. *Science* **249**, 283–286 (1990).

ACKNOWLEDGEMENTS. We thank P. Stanley for providing VSV, J. M. Sheil for the gift of N1 cells and clone 33, B. Bear and R. Angeletti for making the synthetic peptides and for technical assistance, K. Ishikawa and D. S. Lyles for discussing unpublished data, M. Scharff, B. Diamond, B. Birstein for reading the manuscript, S. Joyce and K. Hasenkrug for editorial help, R. Spata for typing, and B. McCue for her involvement. The research was supported by the NIH and the Irvington Institute for Medical Research.

Molecular structure of F-actin and location of surface binding sites

R. A. Milligan, M. Whittaker & D. Safer*

Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, California 92037, USA

* Department of Anatomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Comparisons of three-dimensional maps of vertebrate muscle thin filaments obtained by cryo-electron microscopy and image analysis, reveal the molecular structure of F-actin, the location of the C terminus of the monomer and the positions of the binding sites of tropomyosin, the myosin head and the N-terminal portion of the myosin A1 light chain on the filament. These data provide strong constraints for evaluating models built from the atomic structure of the monomer and the subsequent identification of molecular contacts.

F-ACTIN is a two-stranded helical polymer of actin molecules which, together with the tropomyosin-troponin regulatory complex, forms the thin filaments of skeletal muscle. In muscle, thick (myosin-containing) filaments lie parallel to the thin filaments and the myosin heads with their associated light chains project outwards from the thick filament backbone. The cyclic interaction of the myosin heads (S1) with F-actin results in contraction.

The structures of F-actin, thin filaments and S1-decorated filaments have been investigated several times at low-resolution using negative-staining electron microscopy and image analysis¹⁻⁹. The results from these investigations have not always been in agreement, and conflicting models for the molecular structure of F-actin, the geometry of S1 binding and the position of tropomyosin have been proposed. From some of these data, a consensus model for F-actin has emerged¹⁰. The model describes only the gross features of the filament: monomers are arranged with their long axes almost perpendicular to the filament axis giving an overall filament diameter of 90–100 Å. Comparisons of three-dimensional (3-D) maps of S1-decorated filaments in negative stain have not provided a consensus on the structural details of S1.

Previously we presented results demonstrating that cryo-electron microscopy coupled with image averaging is capable of yielding highly reproducible 3-D maps of muscle filaments¹¹. We described the binding geometry of S1 and, using a statistical approach, determined the location of tropomyosin in the filament. Here, we extend this work and present 3-D maps of F-actin and S1-decorated filaments, which reveal the structure of the actin molecule and the packing and bonding pattern of molecules in filaments. In addition, we have unequivocally localized the C-terminal part of the actin molecule in the filament. Finally, by comparing several 3-D maps, we have determined the locations of the binding sites for tropomyosin, the myosin head (S1) and the N-terminal domain of the A1 myosin light chain on the molecular surface of F-actin. This structural analysis of intact filaments provides the essential framework for constructing an atomic model of F-actin from the high-resolution X-ray structure of the actin monomer¹²⁻¹⁴.

Images and image processing

We used cryo-electron microscopy and helical image processing methods to calculate 3-D maps of F-actin, S1-decorated thin filaments, S1(A2)-decorated actin (F-actin decorated with S1

TABLE 1 Summary of parameters from image analysis

	Datasets*	Phase residual†	Radial scale factor†	Molecules‡
F-actin	19	25° (14–34°)	1.02 (0.98–1.06)	916
Thin filament-S1	34	24° (19–34°)	1.03 (0.98–1.07)	1,339
Actin-S1(A1)	14	22° (16–25°)	0.99 (0.97–1.04)	546
Actin-S1(A2)	24	21° (14–28°)	1.00 (0.97–1.07)	884
Au ₁₁ -Actin-S1(A2)	34	24° (15–32°)	1.01 (0.97–1.07)	1,300

* Number of near- and far-side datasets averaged to give layer-line data shown in Fig. 2.

† Average phase residual (degrees) and average radial scaling factor obtained when individual datasets were fitted to a reference dataset¹⁵. Numbers in parentheses show the range of each parameter for all datasets included in final average. For F-actin, fitting was carried out using the strong peaks on LLs 1, 2 and 5–7. For the decorated filaments, strong peaks on LLs 1, 2 and 4–7 were used.

‡ Total number of molecules (actin or actin-S1) averaged for each 3-D map.

containing light chain A2), S1(A1)-decorated actin and undecagold-labelled actin decorated with S1(A2). We took many images of each preparation and carefully selected the best images in terms of optical quality and filament preservation for computer processing. Individual images were noisy and of low contrast (Fig. 1), but enhancement of the signal was achieved by averaging. Images of the various S1-decorated filament preparations were visually indistinguishable (compare Fig. 1b and c). Typically, layer-lines (LLs) 0, 1, 6 and 7 were strong in diffraction patterns of individual actin filaments and LLs 0–8 were strong in patterns from S1-decorated filaments (Fig. 1, lower panels). The computed transform from each filament gives two datasets: one from the near side and one from the far side of the filament¹⁵. We averaged many of these datasets to provide the layer-lines from which the 3-D maps were calculated (see Table 1). Averaging reduced the noise to a low level and resulted in a strong signal on LLs 0, 1, 2 and 5–8 for F-actin and on LLs 0–8 for the S1-decorated filaments (Fig. 2). All the final datasets extend to a resolution of 25–30 Å radially and ~30 Å axially. Density peaks in the 3-D maps and difference maps are therefore located with a positional accuracy of about ±5 Å. We investigated the differences between the specimens by performing subtractions in reciprocal space with appropriate dataset pairs and then calculating 3-D difference maps. We assessed the significance of peaks in these maps using a real-space, statistically-based method described previously¹¹. Binding sites and amino-acid localizations were determined by displaying the intersection of the molecular surfaces of 3-D maps and difference maps with the molecular surface of F-actin.

F-actin

First we calculated a 3-D map of F-actin to determine the molecular structure of the filament and to serve as the basis for localization of surface sites. Contouring the F-actin map at a level permitting contact between the actin monomers and a tropomyosin molecule of 20 Å diameter in a difference map, results in a maximum filament diameter of 95–100 Å. The map (Fig. 3a) shows that the filament is constructed of actin monomers arranged with their long axes aligned roughly along the genetic helix. As described previously¹¹, the internal density distribution in the monomer suggests that it is composed of two

domains. We refer to them as outer (Ao) and inner (Ai) domains reflecting their proximity to the filament axis. The outer domain is centred ~ 26 Å from the filament axis with the inner domain at a radius of ~ 17 Å. In the map there are connections between adjacent monomers along both the right-hand long pitch helix and the left-hand genetic helix. Long pitch connections join

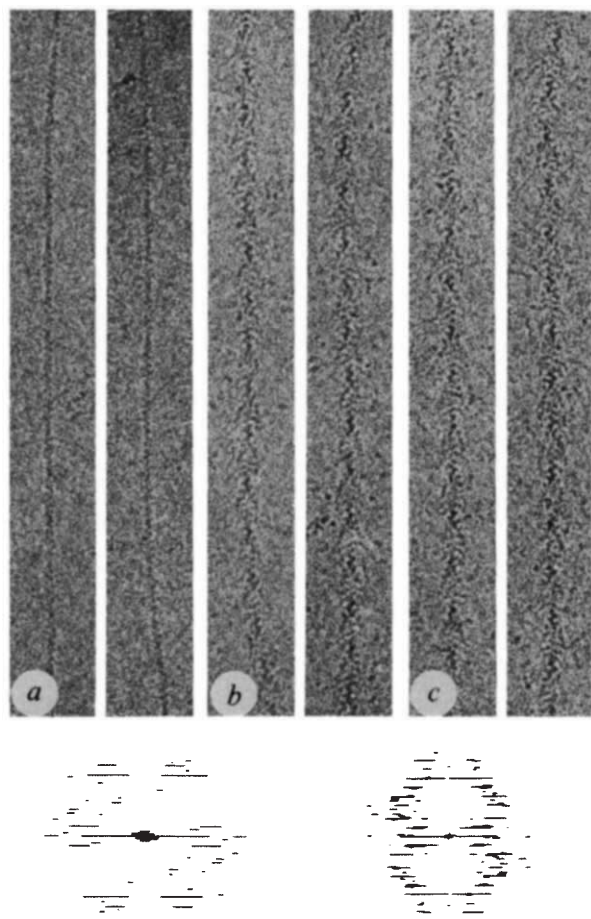


FIG. 1 Representative images and computed diffraction patterns of filaments. *a*, F-actin. A typical diffraction pattern from a straight region is shown at lower left. *b*, S1(A2)-decorated actin. *c*, Undecagold-labelled F-actin decorated with S1(A2), with typical diffraction pattern at lower right.

METHODS. Actin, myosin S1(A1) and S1(A2), the tropomyosin-troponin regulatory complex and synthetic thin filaments were prepared from rabbit muscle by standard procedures³²⁻³⁵. For gold-cluster labelling, monomaleimide undecagold was prepared and coupled to actin as described previously¹⁸. Labelled actin was separated from unincorporated gold by Sephadex S-200 chromatography, polymerized by the addition of salt and then stabilized with a twofold molar excess of phalloidin. Incorporation of undecagold was determined spectrophotometrically and by non-denaturing polyacrylamide gel electrophoresis³⁶ and was found to be virtually quantitative. The specificity of labelling on Cys 374 was confirmed by analysis of tryptic peptides and also by experiments showing that labelling was abolished after specific modification of Cys 374 (D.S. *et al.*, manuscript in preparation). The properties of labelled and unlabelled filaments are similar: both polymerize and depolymerize in response to changes in salt concentration, both form Mg^{2+} paracrystals, and, as shown here, both form complexes with S1 exhibiting identical overall morphology. These observations indicate that the gross structure of the actin molecule has not been altered by labelling with undecagold. Frozen specimens were prepared as described previously¹¹ except that each grid was rinsed with 2-3 drops of a solution containing 10 mM imidazole, pH 7.0, 50 mM KCl, 1 mM $MgCl_2$, 0.5 mM $CaCl_2$, 0.5 mM dithiothreitol just before blotting and freezing. Frozen grids were mounted under liquid nitrogen in a Gatan cryo holder (Model 626) and examined at $-175^\circ C$ in a Philips CM12 microscope equipped with a Gatan anticontaminator and operating at 100 kV. Low dose images of thin vitreous ice spanning holes in the carbon support film were recorded 1.8-2.2 μ under-focus at a nominal magnification of 35,000.

predominantly the inner domains of adjacent monomers, whereas the connections along the genetic helix link the inner domain of one monomer and the outer domain of the next monomer. Both sets of connections lie within ~ 20 Å of the filament axis. There seems to be no major connectivity directly across the helix axis although, because of the limited resolution of this study, we cannot exclude the possibility of minor contacts in this region. Examination of the near-axial regions of the S1-decorated actin maps confirms the findings described above (results not shown). In the F-actin map, each monomer has a slightly concave front face spanning both domains on the outside of the filament. The front, back, top and bottom surfaces of the outer domain seem to be largely accessible whereas on the inner domain, only the portion comprising the front face is accessible.

Tropomyosin binding site

We obtained layer line data contributed by tropomyosin by subtracting the S1-decorated actin data (Fig. 2*c*) from the S1-decorated thin filament data (Fig. 2*b*). Contributions from the troponin complex do not lie on the layer-lines and are therefore not included in the analysis described here. The 3-D map calculated from the difference data is shown superimposed on the F-actin map in Fig. 4*a*. Tropomyosin follows the long pitch helix and lies at a filament radius of ~ 38 Å. Figure 4*b* shows regions where tropomyosin intersects the F-actin surface, that is, the tropomyosin binding site on actin. The elongated binding site lies at the extreme edge of the inner domain of actin leaving most of the outer face of the monomer accessible. It should be noted that this is the site occupied by tropomyosin when both Ca^{2+} and the myosin head are present.

Myosin binding site

The myosin head (S1) interacts with the actin filament in a tangential manner, obscuring the front concave face of the actin monomer (Fig. 3*b*). Below this major site of interaction, there is an additional thin bridge of density extending between the main body of S1 and the top of the outer domain in the adjacent long pitch actin monomer (labelled B in Fig. 3*b*). This bridge lies at a filament radius of ~ 40 Å, and is present in all the S1-decorated filament maps we have calculated.

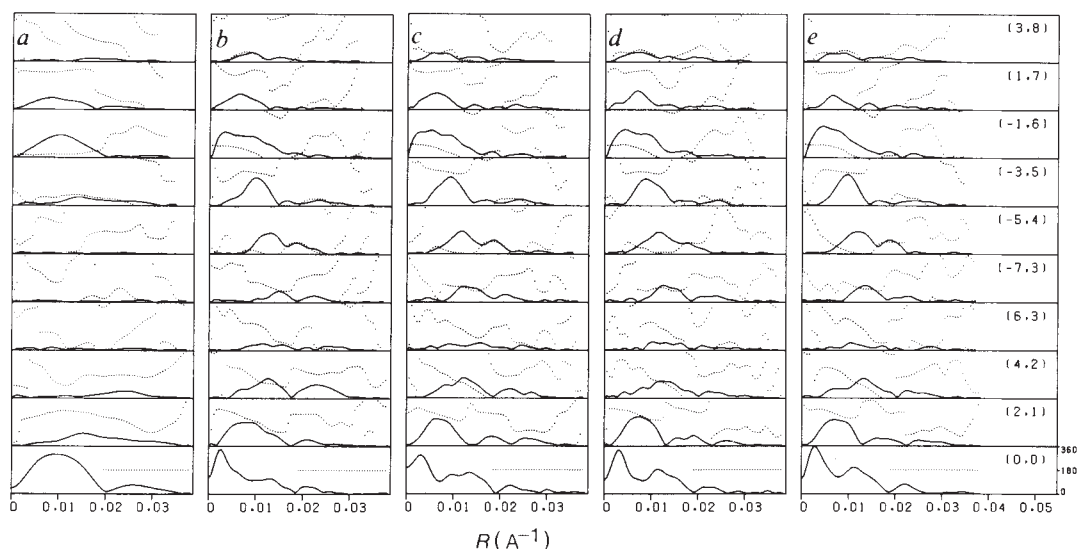
To show the binding site more clearly, the intersection of the F-actin and S1(A2)-decorated actin surfaces is displayed in Fig. 4*c*. As described previously¹¹, the major contact is between S1 and the outer domain of actin but the binding site also extends across the front face of the monomer towards the point of attachment of tropomyosin, suggesting that S1 also binds to the inner domain of actin. The intersection of the two maps also shows clearly the contact point of the bridge on the top surface of the adjacent long pitch actin monomer (arrow, Fig. 4*c*).

Light chains A1 and A2 are interchangeable on the myosin head and differ only in that A1 has an additional N-terminal domain of relative molecular mass 4,200 (M_r 4.2K), which binds to the C-terminal part of actin at residues 360-363^{16,17}. To localize this binding site, we calculated a 3-D map of S1(A1)-decorated actin for comparison with the S1(A2)-decorated actin map. The S1(A1) binding site on actin is similar to the S1(A2) site except that the major contact is larger and extends under the actin monomer (arrow, Fig. 4*d*). A difference map calculated from data obtained by subtracting the two data sets (data in Fig. 2*d* minus data in Fig. 2*c*) shows a single highly significant peak representing the extra N-terminal domain of the A1 light chain, located at a filament radius of ~ 30 Å (Fig. 4*e*). The peak lies on top of the extra region on actin covered by S1(A1). The intersection of this peak with the actin surface (Fig. 4*f*) therefore identifies the location of residues 360-363.

Undecagold labelling of Cys 374 of actin¹⁸ provides independent support for this localization. We determined the position of the gold cluster by comparing maps from gold-labelled and control (unlabelled) S1-decorated actin. The difference map showed a single highly significant peak representing the

FIG. 2 Final averaged layer-line data for F-actin (a), S1-decorated thin filaments (b), S1(A2)-decorated actin (c), S1(A1)-decorated actin (d) and undecagold-labelled actin decorated with S1(A2) (e). Amplitudes are shown as solid lines, phases as dotted lines. Bessel functions (n) and layer line numbers (l) are shown on the final plot as (n, l). All datasets have the same phase origin and amplitudes in b–e are at the same scale.

METHODS. In general, image processing was carried out as described previously¹¹. Briefly, straight regions of filaments were chosen from optimally defocused images that lacked drift and astigmatism. The preservation of these regions was assessed by optical diffraction and the best images converted to optical density arrays for computer processing. Helical processing was carried out in the standard way^{1,2,5,15} assuming 13/6 helical geometry, using a program package written at the MRC Laboratory of Molecular Biology, Cambridge, UK. An integral number of helical repeats—usually 6—were masked, floated into a large array, Fourier transformed, and layer-lines were extracted from the transform. The best position for the helical axis was refined and the near- and far-side datasets from each filament were individually brought to the same phase origin and scaled to a reference dataset. For S1-decorated filaments, the previously published datasets were used as the initial reference¹¹. For undecorated filaments the initial reference was a single near-side dataset. In subsequent rounds of alignment, averages based on parameters obtained in the previous round were used as the reference. Three-dimensional maps were calculated by Fourier-Bessel inversion of the layer-



line data. For reasons discussed elsewhere¹¹, no adjustments were made to the layer line data to compensate for the effects of the contrast transfer function. Conclusions from experiments in which approximate adjustments were carried out, were the same as those presented herein. Undecorated and S1-decorated maps were brought to the same phase origin by a real-space correlation procedure. The densities in common regions of the maps were scaled together and a correlation coefficient was calculated using densities inside the estimated molecular boundary of each map. Relative rotation about the helical axis, translation along the axis, radial scaling of the maps, and polarity were explored to determine all parameters necessary to bring the maps into register, that is, to maximize the correlation coefficient. Data for difference maps were obtained by reciprocal space subtractions and in all cases the significance level of the differences was determined as described previously¹¹.

undecagold cluster at a filament radius of ~ 34 Å located close to the A1 N-terminus difference peak (Fig. 4g). The most likely location for Cys 374 is at the intersection of the gold difference peak and the F-actin surface (Fig. 4h). Thus, Cys 374 lies at a filament radius of ~ 27 Å, at the bottom of the outer domain very close to residues 360–363 (compare Fig. 4f and h).

Discussion

Our analysis of actin filament images provides a description of the molecular structure of F-actin: the long axis of the monomer lies roughly along the genetic helix of the filament. The filament has a maximum diameter of 95–100 Å. The outer and inner domains are centred ~ 26 and ~ 17 Å respectively from the filament axis. In common with others^{7,19}, we find that there are

both longitudinal (long pitch) and diagonal (genetic helix) contacts between monomers. Both sets of bonds lie within ~ 20 Å of the filament axis. The molecular details of F-actin that we have described, are also seen in the near-axial regions of S1-decorated filament maps, attesting to the reliability of our findings.

We used two independent approaches to show that the C-terminal portion of actin lies at the bottom of the outer domain in the monomer. The radial coordinate for Cys 374 is ~ 27 Å, in reasonable agreement with values of 20–25 Å and 35 Å obtained by fluorescence energy transfer^{20,21}. Using the location of Cys 374 as a reference we can also determine the locations of other regions of the sequence in the filament. Crosslinking experiments show that Cys 374 and Lys 191 in adjacent monomers are 12–14 Å apart²². Together with our data, this

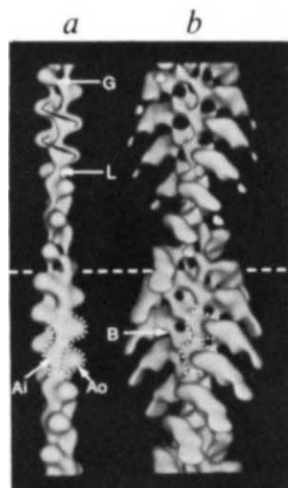


FIG. 3 Three-dimensional maps of F-actin (a) and S1(A2)-decorated F-actin (b) showing the molecular surface (bottom) and, to illustrate connectivities between monomers, the distribution of higher densities in the maps (top). Two adjacent long pitch actin monomers are outlined in the lower half of each map and the outer (Ao) and inner (Ai) domains of the monomer are identified. The left-hand genetic helix is indicated in the upper part of a. The long-pitch connections (L) between inner domains of adjacent monomers, connections joining Ai to Ao along the genetic helix (G) and the bridge from the main body of S1 to the top of Ao (B) are indicated. The maps span two crossovers (720 Å).

METHODS. The molecular boundary of the maps was determined as follows. Tropomyosin difference data were obtained by subtracting S1-decorated actin data (Fig. 2c) from S1-decorated thin filament data (Fig. 2b) and a 3-D map calculated using LLs 0–8. The tropomyosin boundary in the difference map was set at a density level that gave a diameter of ~ 20 Å for the molecule. The boundaries of the F-actin and decorated actin maps were then set at density levels which resulted in contact with tropomyosin.

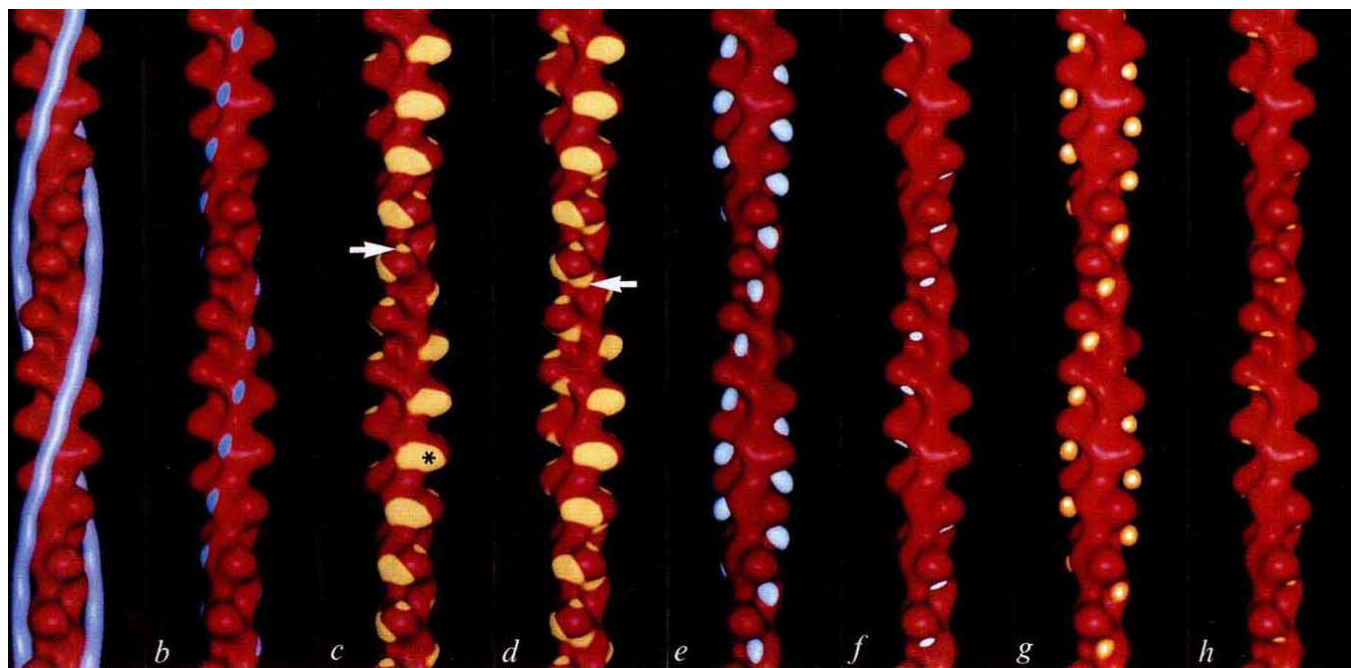


FIG. 4 Difference maps and binding sites displayed on F-actin (red). *a* and *b*, tropomyosin and the tropomyosin binding sites, respectively, are shown in blue. *c* and *d*, The S1(A2) and S1(A1) binding sites, respectively, are shown in yellow. The asterisk marks the location of strongest connectivity between S1 and the outer domain of actin. Arrow on *c* indicates the point of contact of the bridge on the top of Ao. Arrow on *d* indicates the additional area covered by S1(A1) on the bottom of Ao. *e* and *f*, The N-terminal domain of A1 and its binding site, respectively, are shown in white. This binding site includes residues 360–363 of actin^{16,17}. *g*, Location of the undecagold cluster, which is attached to Cys 374 of actin is shown in gold. *h*, The intersection of the F-actin surface and the undecagold peak—the most likely location for Cys 374—is shown in gold. The polarity of the filaments is the same as in the previous figure and in Fig. 3 of Holmes *et al.*²⁸.

METHODS. Difference data for the N-terminal portion of A1 were obtained by subtracting S1(A2)-actin layer lines (Fig. 2c) from S1(A1)-actin layer lines (Fig. 2d). LLs 0–8 of the difference data were used to calculate the map which was contoured at a level which enclosed 100% of the expected volume. Difference data for undecagold were obtained by subtracting S1(A2)-actin layer lines (Fig. 2c) from S1(A2) undecagold-labelled-actin layer lines (Fig. 2e). All layer lines were used to calculate the difference map. Undecagold was displayed as a particle ~20 Å in diameter. The surfaces shown in the difference maps enclose regions of the 3-D maps where the probability that the differences are due to chance is <0.5%¹¹. The binding sites shown in *b*, *c*, *d* and *f* are the intersections of the molecular surfaces of the relevant maps with the molecular surface of F-actin. Maps were displayed using the program RMSVOLUME³⁷.

result limits the possible location of Lys 191 to a region near the genetic helix contact, on the extreme edge of the inner domain near the tropomyosin binding site. Cys 374 and Lys 191 are therefore at opposite ends of the actin molecule. Other crosslinking²³ and ¹H nuclear magnetic resonance studies¹⁷ have shown that the N and C termini of the same monomer are close to each other, suggesting that both termini are in the outer domain. These deductions are borne out in the X-ray crystallographic analysis of actin:DNase I crystals¹⁴. It is noteworthy that the orientation of actin monomers and the radial coordinate of Cys 374 (33 Å) in ribbons in profilin-actin crystals are remarkably similar to the results presented here, lending support to the suggestion that the ribbon and helix states of actin are closely related structurally¹³.

Our findings regarding the position of tropomyosin are in agreement with previous results obtained by negative staining⁶ and cryo-electron microscopy¹¹. In the presence of both Ca²⁺ and the myosin head, tropomyosin is bound to the inner domain of actin close to the genetic helix contact.

The major myosin binding site is located on the lower part of the front face of the outer domain and seems to extend across the inner domain towards the tropomyosin binding site. As well as this major contact with the front face of actin, myosin also seems to have a minor contact on the top of the outer domain of an adjacent long pitch monomer. This is a highly reproducible feature found in all the S1-decorated maps we have calculated. It also seems to be present in the 3-D maps of Vibert and Craig⁵, although in this work the feature was tentatively identified as tropomyosin. Because we have unequivocally localized tropomyosin elsewhere, we can rule out this possibility and conclude

that the feature must represent a direct secondary contact between S1 and actin. Thus, consistent with one interpretation of results obtained from some crosslinking studies^{24,25}, a single myosin head makes contact with two actin monomers.

The binding site for the N-terminal domain of the A1 light chain lies on the bottom of the outer domain, close to both the genetic helix contacts and the long pitch contacts between actin monomers. The binding of A1 in this region may stabilize these contacts as S1(A1), but not S1(A2), promotes actin polymerization *in vitro*²⁶. The demonstration that A1 binds directly to actin also provides a structural basis for the observation that S1(A1) has a higher binding affinity for actin than S1(A2) at low ionic strength²⁷.

Very recently, the atomic structure of the actin:DNase I complex was solved¹⁴ and an atomic model for F-actin has been proposed²⁸. Our data provide strong support for the atomic model of the filament. Specifically, the data and the model are in agreement in terms of filament polarity, monomer orientation, the three dimensional location of residues 360–363 and 374 and, taking into account the different resolutions of the data and the model, the bonding between actin monomers. Our localization of the tropomyosin, myosin and light chain binding sites will now allow an atomic description of these sites on the model. It is anticipated that in the near future, a combination of electron microscopic analysis of filaments such as is presented here, biochemical studies and high-resolution X-ray analysis of crystals of the individual proteins^{12–14,29–31}, will provide a complete understanding of the interactions of actin, myosin and the tropomyosin-troponin regulatory complex at the atomic level. □

Received 3 August; accepted 11 October 1990.

1. Moore, P. B., Huxley, H. E. & DeRosier, D. J. *J. molec. Biol.* **50**, 279–295 (1970).
2. Taylor, K. A. & Amos, L. A. *J. molec. Biol.* **147**, 297–324 (1981).
3. Toyoshima, C. & Wakabayashi, T. *J. Biochem.* **97**, 219–243 (1985).
4. Toyoshima, C. & Wakabayashi, T. *J. Biochem.* **97**, 245–263 (1985).
5. Vibert, P. & Craig, R. *J. molec. Biol.* **157**, 299–319 (1982).
6. O'Brien, E. J., Couch, J., Johnson, G. R. P. & Morris, E. P. in *Actin: Structure and Function in Muscle and Non-Muscle Cells* (eds dosRemedios, C. G. & Barden, J. A.) 3–15 (Academic, Australia, 1983).
7. Trinick, J., Cooper, J., Seymour, J. & Egelman, E. H. *J. Microsc.* **141**, 349–360 (1986).
8. Millonig, R., Sütterlin, R., Engel, A., Pollard, T. D. & Aebi, U. in *Springer Series in Biophysics. Vol. 3, Cytoskeletal and Extracellular Proteins* (eds Aebi, U. & Engel, J.) 51–53 (Springer-Verlag, Berlin, 1989).
9. Wakabayashi, T., Huxley, H. E., Amos, L. A. & Klug, A. *J. molec. Biol.* **93**, 477–497 (1975).
10. Egelman, E. H. *J. Muscle Res. Cell Motil.* **6**, 129–151 (1985).
11. Milligan, R. A. & Flicker, P. F. *J. Cell Biol.* **105**, 29–39 (1987).
12. Sakabe, N. et al. *J. Biochem.* **93**, 299–302 (1983).
13. Schutt, C. E., Lindberg, U., Myslik, J. & Strauss, N. *J. molec. Biol.* **209**, 735–746 (1989).
14. Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. *Nature* **347**, 37–44 (1990).
15. DeRosier, D. J. & Moore, P. B. *J. molec. Biol.* **52**, 355–369 (1970).
16. Sutoh, K. *Biochemistry* **21**, 3654–3661 (1982).
17. Trayer, I. P., Trayer, H. R. & Levine, B. A. *Eur. J. Biochem.* **164**, 259–266 (1987).
18. Safer, D., Bolinger, L. & Leigh, J. S. Jr *J. Inorg. Biochem.* **26**, 77–91 (1986).
19. Erickson, H. P. *J. molec. Biol.* **206**, 465–474 (1989).

20. Kasprzak, A. A., Takashi, R. & Morales, M. F. *Biochemistry* **27**, 4512–4522 (1988).
21. Taylor, D. L., Reidler, K., Spudich, J. A. & Stryer, L. *J. Cell Biol.* **89**, 362–367 (1981).
22. Elzinga, M. & Phelan, J. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6599–6602 (1984).
23. Sutoh, K. & Mabuchi, I. *Biochemistry* **23**, 6757–6761 (1984).
24. Mornet, D., Bonet, A., Audemard, E. & Bonicel, J. *J. Muscle Res. Cell Motil.* **10**, 10–24 (1989).
25. Mornet, D., Bertrand, R., Pantel, P., Audemard, E. & Kassab, R. *Nature* **292**, 301–306 (1981).
26. Chaussepied, P. & Kasprzak, A. A. *Nature* **342**, 950–953 (1989).
27. Chalovich, J. M., Stein, L. A., Greene, L. E. & Eisenberg, E. *Biochemistry* **23**, 4885–4889 (1984).
28. Holmes, K. C., Popp, D., Gebhard, W. & Kabsch, W. *Nature* **347**, 44–49 (1990).
29. Phillips, G. N., Jr, Fillers, J. P. & Cohen, C. *J. molec. Biol.* **192**, 111–131 (1986).
30. Herzberg, O. & James, M. N. G. *J. molec. Biol.* **203**, 761–779 (1988).
31. Rayment, I. & Winkelmann, D. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4378–4380 (1984).
32. Spudich, J. A. & Watt, S. *J. Biol. Chem.* **246**, 4866–4871 (1971).
33. Pardee, J. D. & Spudich, J. A. *Meth. Enzym.* **85**, 164–181 (1982).
34. Weeds, A. G. & Taylor, R. S. *Nature* **257**, 54–56 (1975).
35. Ebashi, S. & Ebashi, F. *J. Biochem.* **55**, 604–613 (1964).
36. Safer, D. *Analyt. Biochem.* **178**, 32–37 (1989).
37. Goodsell, D. S., Mian, I. S. & Olson, A. J. *J. molec. Graphics* **7**, 41–47 (1989).

ACKNOWLEDGEMENTS. We thank D. Goodsell for help with the computer graphics and our colleagues for comments on the manuscript. This work was supported by the National Institutes of Health (R.A.M. and D.S.). R.A.M. is a Pew Scholar in the Biomedical Sciences.

LETTERS TO NATURE

Adaptive optics for array telescopes using neural-network techniques

J. R. P. Angel*, P. Wizinowich*, M. Lloyd-Hart* & D. Sandler†

* Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

† Thermo Electron Technologies Corp., 9550 Distribution Avenue, San Diego, California 92121-2395, USA

IMAGES formed by ground-based telescopes are marred by atmospheric 'seeing'. The plane wavefront from an unresolved star is distorted by continually changing turbulent fluctuations in the air's refractive index. Diffraction-limited performance can in principle be recovered through the methods of adaptive optics, in which the instantaneous wavefront shape is sensed and corrected in real-time by deformable optics that cancel the distortion^{1,2}. The highest resolution will be achieved when this technique is applied to multiple-telescope arrays. For such arrays, the biggest errors caused by seeing at infrared wavelengths are the variations in pathlength and wavefront tilt between array elements. We show here that these errors can be derived by an artificial neural network, given only a pair of simultaneous in-focus and out-of-focus images of a reference star formed at the combined focus of all the array elements. We have optimized a neural network appropriate for 2.2- μ m wavelength imaging at the Multiple Mirror Telescope in Arizona. Corrections made by moving the beam-combining mirrors will largely recover the diffraction-limited profile, with a resolution of 0.06 arcsec.

The imaging properties of an optimum phased array are the same as for a single large telescope masked with apertures in the positions of the array elements. In the absence of atmospheric distortion the image of each star will have a central peak of the same width as that of the unmasked diffraction-limited telescope. The effect of atmospheric seeing is to degrade the instantaneous image into a pattern of random speckles. If the array elements are not too large, then the atmospheric wavefront errors can be adequately corrected in the beam-combining optics, without the need for correcting the wavefront distortion across individual elements.

The degree to which the diffraction-limited image profile of an array can be restored by the piston and tilt correction of individual elements can be calculated from atmospheric turbulence theory. Kolmogorov³ predicted a spatial correlation of turbulence which is in accord with measurements. Fried⁴ showed that removal of the best-fit plane to a circular segment of

Kolmogorov distorted wavefront resulted in a residual mean square phase error of

$$\Delta = 0.130(D/r_0)^{5/3} \quad (1)$$

where D is the segment diameter and r_0 , a function of wavelength and seeing quality, is Fried's turbulence distance scale. If each array element is corrected with the tip, tilt and piston motion of a beam-combining mirror, this error is also appropriate for the entire array.

A measure of the quality of a near-diffraction-limited image is its Strehl ratio, the ratio of the peak intensity to that of the diffraction-limited image⁵. For small phase errors, $\Delta < 1 \text{ rad}^2$ the Strehl ratio can be approximated by⁴

$$\text{SR} \approx e^{-\Delta} \quad (2)$$

Strehl ratios less than unity correspond to loss of contrast rather than resolution, so most of the advantage for astronomical imaging and spectroscopy is maintained for $\text{SR} > 0.5$, corresponding to $D/r_0 < 2.7$.

At visible wavelengths ($\lambda = 0.5 \mu\text{m}$), r_0 is typically 15 cm at good sites, and the element size satisfying the above condition is rather small. The situation is more attractive in the infrared because r_0 scales as $\lambda^{6/5}$. This also implies slower correction rates because the timescale for fluctuations is $\tau \approx r_0/v$, where v is the wind speed of the turbulent air. At 2.2- μm wavelength, $r_0 \approx 1 \text{ m}$ and $\tau \approx 50 \text{ ms}$, and array of elements of 2-m diameter can be used. Other advantages of the infrared wavelength include a higher density of field stars for wavefront sensing and a larger isoplanatic patch (the area of sky over which the wavefront corrections deduced from a central reference star are valid).

A practical method to sense the piston and tilt errors is required. Measurements need to be completed in $\sim 20 \text{ ms}$, if the wavefront over 2-m elements is to remain effectively frozen until the correction is applied. Although no array has yet been adaptively phased, techniques have been developed to bring them into phase for the time-averaged wavefront. Piston errors between pairs of array elements are typically removed by adjusting path differences until Young's fringes are observed in overlapping star images in white light.

For an array of n elements, it could thus be possible to determine the piston errors from the instantaneous fringe position in n pairs of overlapping images and the individual tilt errors from the centroids of n separate images, but the resulting system would be complex and inefficient.

The common methods used for sensing across single mirrors are unfortunately not applicable for an array. Shearing interferometry and the Hartmann-Shack method measure the local wavefront slope at points in the pupil plane⁶. A method by

Roddier measures the local curvature instead⁷. In both cases an integration across the pupil must be made to obtain the shape of the wavefront. For an array, where the wavefront is not continuous, these methods cannot be used.

A method recently demonstrated for single apertures (D.S. *et al.*, manuscript in preparation) offers hope for rapid sensing of the full wavefront of an array. The method makes use of the fact that a wavefront can be reconstructed entirely from the information contained in an in-focus and out-of-focus pair of images, taken simultaneously. Opticians routinely use far-field images to recognize low-order aberrations such as astigmatism or coma and can determine the sign of the aberration by going through the focus. There is presently a massive effort to characterize precisely the wavefront of the Hubble telescope from sets of defocused images. Sandler and Barrett have found that a powerful wavefront sensor can be made by training an artificial neural net to recognize aberrations from the image pairs. The low-order aberrations (the first 18 Zernike coefficients) are determinable from images formed by passing a Kolmogorov-distorted wavefront through a single circular aperture. Previous discussions of neural networks in adaptive optics have largely concentrated on using their parallel processing architecture to rapidly solve existing algorithms^{8,9}. Sandler *et al.*, however, place on the neural net the entire burden for relating the image to the wavefront.

Because far-field images result from the interference of photons from every point in the pupil, the method should not be limited to continuous apertures. The possibility of determining errors for individual elements in an array of telescopes was suggested by Paxman and Fienup¹⁰. They showed how a pair of in-focus and out-of-focus images and phase-retrieval techniques could be used to recover small piston errors between array elements. Because of the extreme nonlinearities in the image formation, however, analytical techniques are very difficult for the case of general wavefront aberration.

The artificial net of Sandler *et al.* was trained on images with properties largely determined by geometric optics. By contrast, the effects of interference are critical to finding the phase errors across an array, and it was not clear how well a neural net would work, although their nonlinear structure seemed well suited to the task. In fact, this approach is extraordinarily powerful, as shown by the following modelling of its application to the Multiple Mirror Telescope (MMT).

The six MMT¹¹ mirrors are arranged hexagonally on a radius of 2.52 m (Fig. 1a). Operating as a co-phased array¹², the image resolution is equivalent to a single filled aperture of 7.64-m diameter¹³. The 1.83-m element size is ideally suited for the near infrared, where the median MMT seeing corresponds to $r_0 = 0.79$ m at $2.2 \mu\text{m}$ ^{14,15}. For the purpose of modelling, we assumed Kolmogorov turbulence with an r_0 of 1 m at $2.2 \mu\text{m}$, typical of

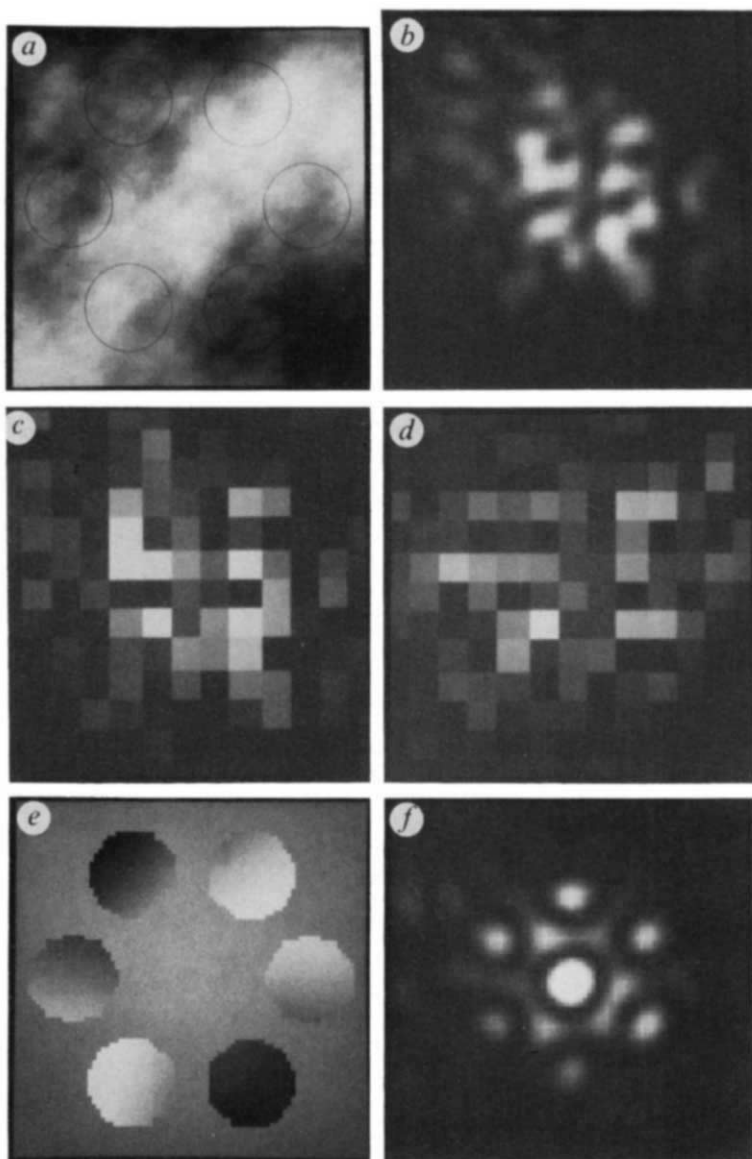


FIG. 1 Passage of a single speckle image, at $2.2\text{-}\mu\text{m}$ wavelength, through a trained neural network. *a*, The initial simulated wavefront with mean tilt removed (the 6 MMT apertures are superposed); *b*, the corresponding in-focus image; *c*, the binned in-focus and *d*, out-of-focus images input to the net; *e*, the net derived wavefront for each aperture; *f*, the image corrected by removing the net derived wavefront tilt and piston. The grey scale in *a* and *e* represents a total phase difference of 7.4 rad. The image scale for *b* and *f* is 0.90 arcsec along a side and for *c* and *d* is 0.65 arcsec.

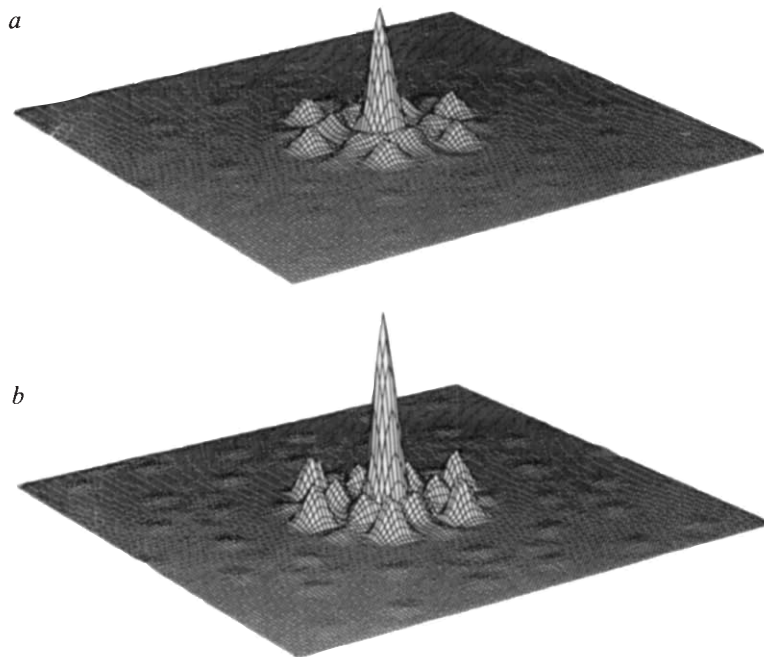


FIG. 2 *a*, A long-exposure image, $SR=0.66$, at $2.2\text{-}\mu\text{m}$ wavelength obtained by adding 500 simulated speckle images. The speckle images input to the net were corrected by removing the tilt and piston for each mirror as determined by the neural-net wavefront sensor. This should be compared with *b*, the theoretical unaberrated image profile for the co-phased MMT. The scale for both plots is 1.3 arcsec along each side.

good MMT seeing. The mean wavefront tilt was removed in the images shown to the neural net. Such correction is useful to reduce the required number of detector pixels, and could be performed with a fast steering mirror and a small detector array. We used a Fraunhofer-diffraction calculation to trace the wavefront from the six mirrors to detector arrays at the in-focus and out-of-focus image planes.

We trained a perceptron class of supervised neural net¹⁶. The in-focus and out-of-focus images are superposed on two detector arrays, where each pixel represents an input node to the net. The input values are the normalized image intensities I_i at each pixel. A single layer of hidden nodes is used. At each hidden node a weighted sum over all the input pixel values is formed. This sum is then passed through a sigmoid function to form the node output

$$o_j = \frac{1}{1 + e^{-(\sum_i w_{ji} I_i - \theta_j)}} \quad (3)$$

which turns on as the sum reaches a certain threshold. The indices i, j and k refer to the input, hidden and output nodes, respectively, the W are the connecting weights, and the θ are constant offsets. There are 18 output nodes, whose signals are weighted sums over the hidden-node outputs. These outputs

$$o_k = \sum_j W_{jk} o_j \quad (4)$$

are the x tilt, y tilt and piston adjustment for each of the six MMT mirrors.

The goal of net training is to determine optimum values for the weights connecting the input, hidden and output layers, and the sigmoid offsets. Initially all are set to small random values. The net is then shown a large set of image pairs calculated from random Kolmogorov-distorted wavefronts, for a given value of r_0 . For each image pair the net outputs are compared to the known wavefront tilts and pistons, and the net weights and offsets are adjusted to bring the outputs into better agreement. The formal back-propagation algorithm, based on the least-squares method, was derived by Rumelhart *et al.*¹⁷.

The network for $2.2\text{-}\mu\text{m}$ images consists of 338 input nodes (two 13×13 arrays) and 150 hidden nodes. Training was accomplished with 2.5×10^5 simulated image pairs. The ability of the trained net to determine wavefront errors is illustrated in Fig. 1. A typical simulated wavefront is shown in Fig. 1*a*, with super-

posed circles representing the six MMT mirrors. Figure 1*b* shows the resultant in-focus speckle image. The corresponding in-focus and out-of-focus images, in a binned 13×13 format for input to the neural net, are shown in Fig. 1*c, d*. From these images alone the net derived the piston and tilt errors shown in Fig. 1*e*, which can be seen to have the same character as the original wavefront. After applying these corrections the mean square error is reduced from 2.27 to 0.41 rad^2 . The corresponding corrected image, Fig. 1*f*, is quite similar to that of the unaberrated array.

We modelled the image that would result from continuous on-line correction by adding together 500 individual images corrected by the neural net. The result is an integrated image that is nearly diffraction-limited, Fig. 2*a*, with a Strehl ratio of 0.66 (equations (1) and (2) predict $SR=0.70$). The energy lost from the perfect image pattern, whose central spike has a full width at half maximum (FWHM) of 0.06 arcsec , is spread into the expected^{1,2,6} low-level seeing disk, with a $FWHM \approx (\lambda/r_0) = 0.45\text{ arcsec}$. A comparison with the theoretical point-spread function of the co-phased MMT, shown in Fig. 2*b*, illustrates the success of this neural-net wavefront sensor.

We have made a preliminary analysis of the effect of photon and detector noise. K-band ($2.2\text{-}\mu\text{m}$ wavelength) detectors with a read-out noise of 10 electrons now seem within reach (G. Rieke, personal communication). Simulations with this much noise show that the neural-network approach will work well at the MMT for reference stars as faint as K magnitude 10. The isoplanatic angle at $2.2\text{ }\mu\text{m}$ is about one arc minute and the average density of such reference stars is 20 per square degree¹⁸. Thus a total of some hundred square degrees of sky, in one-arcminute patches, is amenable to high-resolution imaging by this method.

The above example points to the wider application of this neural-net approach, even beyond atmospheric compensation. In particular, this method would be well suited to the sensing of the alignment errors of the telescope arrays being planned for high-resolution imaging in space. $\square$

Received 5 July; accepted 3 October 1990.

1. Beckers, J. M. *ESO Conf. Very Large Telescopes and Their Instrumentation* Vol. 2 (ed. Ulrich, M.-H.) 627–638 (European Southern Observatory, Garching bei München, 1988).
2. Rousset, G. *et al. Astr. Astrophys.* **230**, L29–L32 (1990).
3. Kolmogorov, A. *Turbulence* (eds Friedlander, S. K. & Topper, L.) 151–155 (Interscience, New York, 1961).

4. Fried, D. L. *J. opt. Soc. Am.* **55**, 1427–1435 (1965).
5. O'Neill, E. L. *Introduction to Statistical Optics*, 87–88 (Addison-Wesley, Reading, Mass., 1963).
6. Hardy, J. W. *Proc. Soc. Photo-instrum. Engrs* **1114**, 2–13 (1989).
7. Roddier, F. *Appl. Opt.* **27**, 1223–1225 (1988).
8. Smithson, R. C. *Lockheed Horizon* **28**, 24–31 (1990).
9. Merkle, F. *ESO Conf. Very Large Telescopes and Their Instrumentation Vol. 2* (ed. Ulrich, M.-H.) 656 (European Southern Observatory, Garching bei München, 1988).
10. Paxman, R. G. & Fienup, J. R. *J. opt. Soc. Am.* **A5**, 914–923 (1988).
11. Hege, E. K., Beckers, J. M., Strittmatter, P. A. & McCarthy, D. W. *Appl. Opt.* **24**, 2565–2576 (1985).
12. McCarthy, D. W., McLeod, B. A. & Barlow, D. *Proc. Soc. Photo-instrum. Engrs* **1237**, 496–507 (1990).
13. Born, M. & Wolf, E. *Principles of Optics* 6th edn, 416 (Pergamon, Oxford, 1984).
14. Chaffee, F. H. & Cromwell, R. H. *Proc. Soc. Photo-instrum. Engrs* **1236**, 13–17 (1990).
15. Cromwell, R. H., Haemmerle, V. R. & Woolf, N. J. *Proc. Soc. Photo-instrum. Engrs* **1236**, 520–529 (1990).
16. Lippmann, R. P. *IEEE Acoustics, Speech, and Signal Processing Mag.*, 4–22 (April, 1987).
17. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. *Parallel Distributed Processing Vol. 1* (eds Rumelhart, D. E. & McClelland, J. L.) 318–362 (MIT Press, 1986).
18. Allen, C. W. *Astrophysical Quantities* 3rd edn, 244 (Athlone, London, 1973).

ACKNOWLEDGEMENTS. We thank R. Fugate for encouragement. This work was supported in part by the US Air Force Weapons Laboratory.

An upper limit to violations of the Pauli exclusion principle

D. Kekez*, A. Ljubičić* & B. A. Logan†

*Ruder Bošković Institute, PO Box 1016, 41001 Zagreb, Croatia, Yugoslavia

†Ottawa-Carleton Institute of Physics, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5

THE Pauli exclusion principle was formulated to explain the regularities of both the Periodic Table and atomic spectra¹, and has a crucial role in our understanding of the structure and properties of atoms and nuclei. Despite its enormous success the absolute validity of this principle is still open to question and some recent theoretical analyses have suggested that small violations are possible. This has inspired several experimental tests of its validity. Here we consider the possibility of a β -decay process that would violate the Pauli exclusion principle and use existing data from an underground laboratory to place a severe limit on the decay probability, and on possible violations of the Pauli exclusion principle.

In an extension to the field-theory analysis by Ignatiev and Kuzmin^{2,3} of a single oscillator, Greenberg and Mohapatra⁴ developed a theory that allows for small violations of the Pauli exclusion principle (PEP). The analysis, to some extent a development of some ideas due to Green⁵, involves trilinear commutation relations that include a small parameter $\beta \ll 1$; $\beta^2/2$ is a measure of the violation of the PEP and is defined as the relative probability of finding two fermions in the same state. Govorkov⁶ showed that their theory predicted states with negative probabilities, however, and it was concluded that it would not be possible to construct a field theory for small PEP violations⁷. The possibility of 'infinite statistics' was not treated by Govorkov; in such statistics no assumption is made restricting the number of particles in symmetric or antisymmetric states. These statistics have been studied^{8–10}, and Mohapatra^{9,10} has suggested that it may be possible to construct a new kind of field theory in which small violations of the PEP could be allowed.

Experimental tests of the validity of the PEP are of considerable interest and the status of the current and proposed activity has been reviewed recently^{11–13}. Novikov *et al.*¹⁴ investigated PEP violations by using accelerator mass spectrometry to search for anomalous atoms with three electrons in the K shell and obtained a limit of $\beta^2/2 < 2 \times 10^{-21}$. Ramberg and Snow¹⁵ placed an upper limit on $\beta^2/2$ of 1.6×10^{-26} by searching for K-shell X-ray emission when a current is passed through a copper strip; the X-rays would be emitted if the 'new' electrons cascade down to an already occupied K shell. Recently we investigated the non-paulian electron-capture decay of ⁷¹Ge and deduced a limit

of $\beta^2/2 \leq 3 \times 10^{-12}$ for neutrons¹⁶. We also considered the possibility of non-paulian decays in stable nuclei through β^- (β^+) decays of higher-shell neutrons (protons) directly into the lower proton (neutron) shell. It is well known that if the hamiltonian of a system is permutationally invariant, transitions to a filled shell are forbidden for reasons not related to the PEP. But these restrictions are limited to systems in which the number of a given fermion does not change. In the case of β^- (β^+) decay both the number of protons and neutrons changes and the general restrictions for transitions to a filled shell do not apply. In particular, we studied the non-paulian decays of ¹²⁷I into the non-paulian nuclei ¹²⁷Xe or ¹²⁷Te. We searched for β^- (β^+) particles with a large NaI(Tl) detector and obtained a limit of $\beta^2/2 \leq 2 \times 10^{-26}$.

A much more severe limit on $\beta^2/2$ can be deduced using the existing data from underground laboratories. The liquid scintillation detector^{17,18} (LSD) of the Mont Blanc collaboration is particularly suitable. It has been in operation since 1984 in a cavity in a tunnel linking France and Italy, about 1,700 m below the surface (5,200 m water-equivalent). The LSD contains 90 tons of liquid scintillator in 72 stainless steel containers, each with a volume of 1.5 m³. The density of the liquid scintillator C_nH_{2n+2} ($\bar{n} = 10$) is 0.8 g cm⁻³ and there are 5×10^{28} carbon atoms in each container.

In ¹²C the 2p shell is ~ 20 MeV above the 1s shell. The most probable non-paulian transition is expected to be the first-forbidden $2p \rightarrow 1s$ β^- (β^+) transition to ¹²N(¹²B). For a conventional decay the log *ft* values for such first forbidden transitions range from 6 to 9; a conservative value of 9 corresponds to a half-life of ~ 300 s. In each container the Pauli-allowed rate N_p for a β -decay to a 1s shell with a vacancy is then 1.7×10^{26} s⁻¹.

During a 75-day run only two events with energies ≥ 12 MeV were recorded in the inner 7.2 tons of the LSD¹⁹. In 1988 more shielding was added and this new configuration (LSD-2) has reduced background at lower energies²⁰; the background for $E \geq 12$ MeV, however, did not decrease. The counting rate of 3.1×10^{-7} s⁻¹ for $E \geq 12$ MeV can be used as a limit on N_{PV} , the rate of the PEP-violating decays of ¹²C to the occupied shell of ¹²N(¹²B). The corresponding N_p rate for 7.2 tons is 0.97×10^{27} s⁻¹. We expect about half of the non-paulian decays to emit β -particles with energies above the LSD bias level of 12 MeV. In our case $\beta^2/2 = \lambda_{PV}/\lambda_P \leq 2N_{PV}/N_p = 6.5 \times 10^{-34}$ where λ_{PV} and λ_P are the decay constants.

This is a very severe limit on $\beta^2/2$ and is many orders of magnitude more restrictive than the other available experimental limits. The decay of ¹²C is probably not an ideal candidate for studying PEP violations and it may be possible to improve the limit by studying the non-paulian decay of other nuclei in underground laboratories. □

Received 9 July; accepted 2 October 1990.

1. Pauli, W. *Z. Phys.* **31**, 765–783 (1925).
2. Ignatiev, A. Yu. & Kuzmin, V. A. *Yadernaya Fiz.* **46**, 786–790 (1987).
3. Ignatiev, A. Yu. & Kuzmin, V. A. *Soviet J. nucl. Phys.* **46**, 444–446 (1987).
4. Greenberg, O. W. & Mohapatra, R. N. *Phys. Rev. Lett.* **59**, 2507–2510 (1987).
5. Green, H. S. *Phys. Rev.* **90**, 270–273 (1953).
6. Govorkov, A. B. *Phys. Lett.* **A137**, 7–10 (1989).
7. Greenberg, O. W. & Mohapatra, R. N. *Phys. Rev. Lett.* **62**, 712–714 (1989).
8. Greenberg, O. W. *Phys. Rev. Lett.* **64**, 705–708 (1990).
9. Mohapatra, R. N. *Physics Preprint* 90–120 (Univ. of Maryland, 1990).
10. Mohapatra, R. N. *Bull. Am. phys. Soc.* **35**, 981 (1990).
11. Okun, L. B. *Festi-Val-Festschrift for Val Telegdi* (ed. Winter, K.) 201–211 (Elsevier, Amsterdam, 1988).
12. Okun, L. B. *Comments nucl. Particle Phys.* **19**, 99–115 (1989).
13. Greenberg, O. W. & Mohapatra, R. N. *Phys. Rev. D* **39**, 2032–2038 (1989).
14. Novikov, V. M. *et al. Phys. Lett.* **B240**, 227–231 (1990).
15. Ramberg, E. & Snow, G. A. *Phys. Lett.* **B238**, 438–441 (1990).
16. Kekez, D., Ljubičić, A., Kaucić, S. & Logan, B. A. *Z. Phys.* (in the press).
17. Badino, G. *et al. Nuovo Cim.* **C7**, 573–590 (1984).
18. Aglietta, M. *et al. Nuovo Cim.* **C9**, 185–195 (1986).
19. Aglietta, M. *et al. Proc. Int. Symp. on Weak and Electromagnetic Interactions in Nuclei* (ed. Klapdor, H. V.) 741–744 (Springer, Berlin, 1986).
20. Aglietta, M. *et al. Nucl. Phys.* **B34**, 204–209 (1990).

ACKNOWLEDGEMENTS. We thank D. Dehler for discussions and O. Saavedra for a personal communication. B.A.L. thanks the Natural Sciences and Engineering Research Council of Canada for support.

Direct femtosecond mapping of trajectories in a chemical reaction

A. Mokhtari, P. Cong, J. L. Herek & A. H. Zewail

Arthur Amos Noyes Laboratory of Chemical Physics,
California Institute of Technology, Pasadena, California 91125, USA

IN chemical reactions, the dynamics of the transition from reagents to products can be described by the trajectories of particles (or rigorously, of quantum mechanical wave packets) moving on a potential-energy surface. Here we use femtosecond pulsed laser techniques to follow directly the evolution in space and time of such trajectories during the breakage of a chemical bond in the dissociation of sodium iodide. The bond breakage can be described in terms of the time evolution of a single reaction coordinate, the internuclear separation. As the velocities of the separating fragments are typically of the order of a kilometre per second, a time resolution of a few tens of femtoseconds is required^{1,2} to view the motions on a molecular distance scale of less than an ångström. The resolution obtained here permits the direct visualization of the wave packet's motion and provides snapshots of the trajectories along the reaction coordinate.

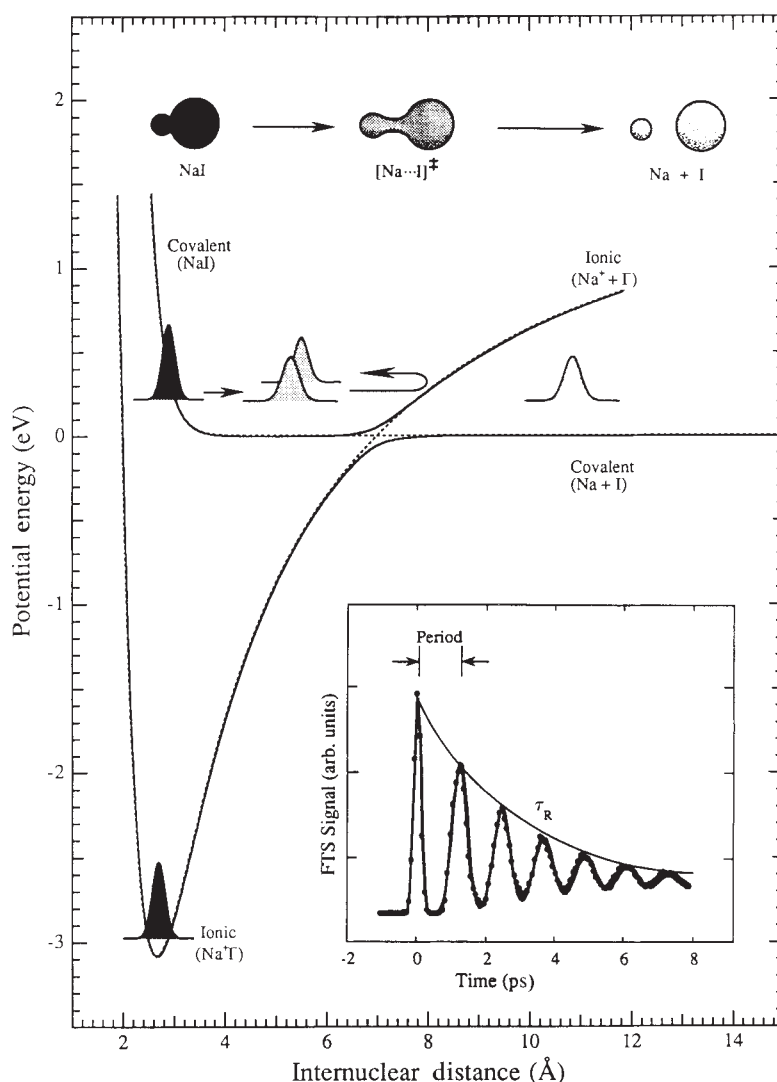
The potential-energy surface (PES) for the NaI reaction (Fig. 1) involves two channels; one describes the covalent character of the bond ($\text{Na} + \text{I}$) and the other the ionic character ($\text{Na}^+ + \text{I}^-$).

The two potential curves intersect at an internuclear separation R of 7 Å. When NaI is promoted with a femtosecond (pump) pulse to the covalent curve at $R \approx 2.7$ Å, the wavefunction (or, more accurately, the wave packet) moves towards the ionic curve and then returns with a period determined by the shape of the PES and the total energy in the bond^{3,4}. Using a femtosecond probe pulse, we can monitor the oscillation of this activated complex⁵, $[\text{Na} \cdots \text{I}]^\ddagger$ as it decays, on a picosecond timescale, to give Na and I (Fig. 1 inset).

Figure 2 shows trajectories in the R - t domain. Both classical newtonian mechanics^{3,4,6} and quantum mechanics⁷ calculations have been performed; here we consider the former approach to illustrate the methodology. As shown in Fig. 2, the wave packet changes direction as R and t change, starting from the initial time zero at $R \approx 2.7$ Å. After half a period the packet has reached the ionic turning point, and after a complete period it has returned to the covalent turning point. At or near the ionic-covalent crossing, the packet splits; some continues to move in the direction of increasing R , becoming Na and I, and the rest remains as $[\text{Na} \cdots \text{I}]^\ddagger$, oscillating between the covalent and ionic points. (See also Fig. 1.)

To observe the motion of the trajectories, a window in $R(t)$ must be opened with sufficient distance resolution ΔR to allow us to view the directionality of the packet at different times. If the window is stepped continuously along R , then we can map out the motion and characteristics of the PES. For a given resolution, the temporal motion represents a 'snapshot' of the

FIG. 1 Schematic diagram of the wave-packet motion in the dissociation reaction of NaI. The femtosecond pump pulse 'transfers' the ground-state wavefunction to the adiabatic potential formed by the avoided crossing of the ionic and covalent curves. The wave packet then moves back and forth, leaking a little ($\sim 10\%$) to $\text{Na} + \text{I}$ every time it approaches the crossing region. The probe femtosecond pulse monitors this oscillatory motion and the decay (τ_R), as shown in the insert.



PES at this particular R . For example, a window at $R = 3 \text{ \AA}$ will give the snapshot shown at the bottom of Fig. 2, which is basically the oscillatory motion displayed in Fig. 1. On the other hand, if a snapshot is taken at $R = 4.5 \text{ \AA}$, the oscillatory motion displays a splitting because at this distance the packet is going in and out of the probing ΔR window. At longer R the splitting increases in time, because the distance the trajectory travels is longer. The observation of such splittings would indicate two important points. First, it would mean that the window opened in R is sufficiently narrow⁸ to resolve the nuclear motion in and out of that region of the PES. Second, the splitting gives a time (at each R) that is directly related to the clocking time, defined⁹ by the distance travelled from time zero to the probing point on the PES.

The experiments required pump and probe pulses tunable between 310 nm and 390 nm, and between 590 nm and 700 nm, respectively. Briefly, a colliding-pulse mode-locked ring dye laser, generating 50-fs pulses around 620 nm, was amplified in a Nd:YAG-pumped four-stage dye amplifier. The output pulses from the amplifier were temporally recompressed in a sequence of four high-refractive-index glass prisms. The tunability on the probe arm was achieved by generating a white-light continuum, and the same procedure was used on the pump arm, except that the light was further amplified in a flowing cell of the appropriate laser dye (pumped by 532-nm pulses from the Nd:YAG source). We used nonlinear doubling and mixing techniques for conversion to ultraviolet. The pump and probe beams were focused in a NaI reaction chamber and the laser-induced fluorescence was collected perpendicular to the propagation direction of the pump and probe beams. The time delay between the pump and

probe was controlled by a high-precision Michelson interferometer and the experiments were controlled by a computer. The range of R probed is shown in Fig. 3.

Figure 4 shows our experimental observations for the NaI reaction. The packet was prepared by a femtosecond pulse at a wavelength λ_1 . To probe at different R , we used another femtosecond pulse of different wavelength λ_2 . For a given total energy in the bond, determined by λ_1 , both the time and the λ_2 can be changed systematically to obtain the snapshots at different R . In Fig. 4, we show three snapshots for $\lambda_1 = 391 \text{ nm}$ as examples of the many snapshots we took. The evolution of the splitting as λ_2 decreases (that is, as R increases) is evident. The splitting at $\lambda_2 = 620 \text{ nm}$, for example, is 350 fs. This gives a clocking time for the packet to reach the window on the PES of 175 fs. The actual distance being probed is $5.6 \text{ \AA}$, as determined by the velocity of recoil. The window resolution is very appropriate for the dynamical motion of chemical reactions: the temporal splitting observed gives an experimental $\Delta R \approx 0.5 \text{ \AA}$ on the PES. The potential for these types of reactions is quite flat at intermediate R (see Fig. 1), and therefore even better resolution may be achieved in other reactions.

For NaI dissociation, the probe at λ_2 excites the wave packet to an upper surface corresponding to the $\text{Na}^* + \text{I}$ channel. The relationship between λ_2 and R is direct if this upper surface is flat at longer R , as discussed in refs 3 and 4. We have determined that there is a well (depth $\approx 1,500 \text{ cm}^{-1}$, which is deeper than the well reported in ref. 10) in the region probed, which allows us to look at the different regions of the PES shown in Fig. 3. For a given λ_2 , we can determine characteristics of the covalent curve by changing λ_1 —direct clocking on the lower surface.

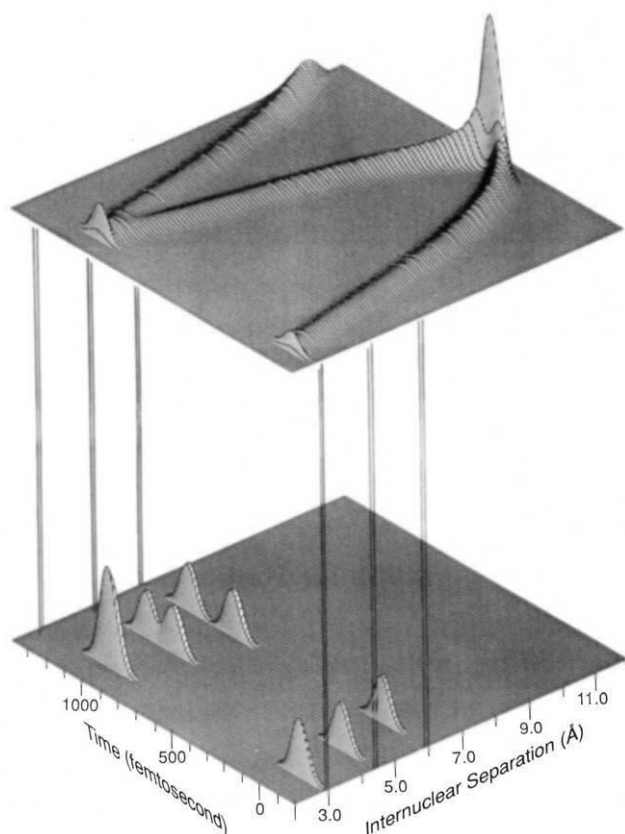


FIG. 2 Classical mechanics simulation of the wave-packet motion with a femtosecond pump pulse at 320 nm. The upper plot shows the wave-packet motion in the two-dimensional space of time and internuclear separation. For simplicity, we do not show the leakage at the crossing point. The lower plot shows the snapshots of the wave-packet motion at different internuclear separations. Note the evolution of the splitting.

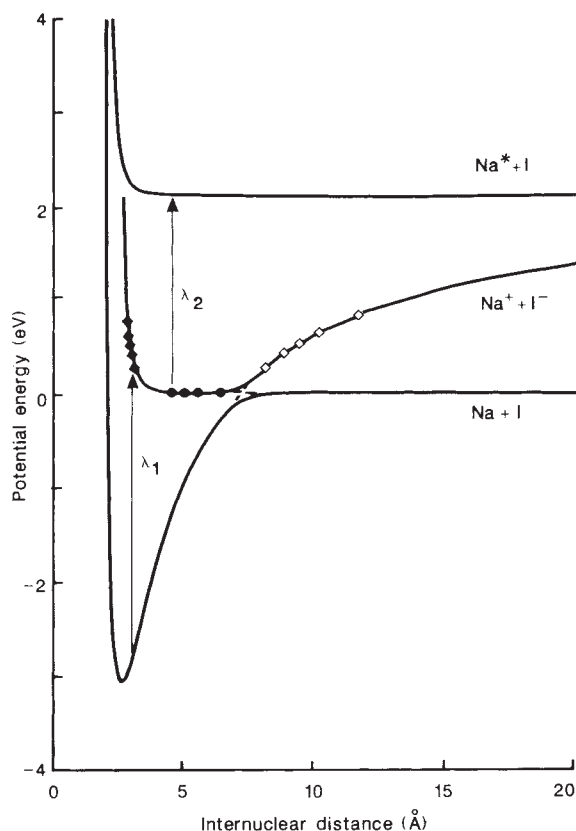


FIG. 3 The range of internuclear distances probed in the experiments. The filled diamonds represent the range that can be reached by pump (λ_1) pulses, the filled circles show the internuclear separations accessed by the probe (λ_2) pulses, and the open diamonds show the points that can be obtained from the oscillation periods. The potentials drawn are taken from ref. 2. We plan to use the snapshots results to deduce more accurate potentials⁸.

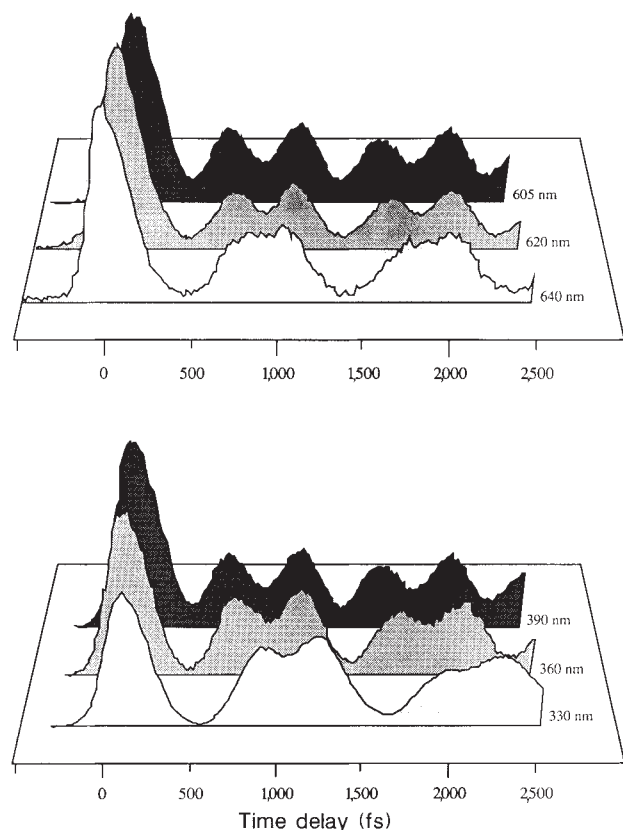


FIG. 4 Experimental femtosecond snapshots of NaI reaction. The upper panel shows the snapshots taken at different probe wavelengths for a wave packet prepared by exciting the NaI molecule at 391 nm (fixed total energy in the bond). The lower panel shows the snapshots taken at a constant λ_2 of 605 nm, while varying the original position (R_0) of the wave packet (different total energies).

Changing the total energy $E_1 = \frac{1}{2}\mu v^2$ and repeating the probing gives the motion on different parts of the PES. The reduced mass μ is constant and by changing λ_1 or E_1 one changes the velocity of recoil v . As E_1 is decreased, we expect the velocity to decrease and the distance travelled by the wave packet, to a given R , to also decrease. Depending on the PES, these two effects determine whether or not the splitting will increase or decrease with E_1 . Figure 4 shows three snapshots at different λ_1 . We observe an increase in the splitting as E_1 is lowered. We also observe a decrease in the period. The latter is consistent with the potential 'narrowing' in R at lower E_1 . The former indicates that the covalent side of the potential is very repulsive at short R and is dominated by velocity (not distance) effects, as shown in Fig. 3.

Theoretically, the use of classical mechanics may be sufficient to describe the global motion on the PES. In fact, the results of our experiments and the classical mechanics calculations in Fig. 2 are in agreement. We have, however, done quantum mechanical calculations to justify the agreement between the classical trajectory $R(t)$ and the quantum mechanical expectation value $\langle R(t) \rangle$ (ref. 11). Quantum mechanical calculations at three pump wavelengths, 310 nm, 350 nm and 390 nm indicate that the results agree with those for classical mechanics. The only discrepancy occurs close to the turning point on the steep repulsive branch—this arises because the average position of the wave packet does not reach the classical turning point. Rigorous quantum mechanical calculations of the wave-packet propagation¹² have shown evidence for the splitting; the interpretation did not address the clocking and mapping-of-trajectories aspect of the problem. For NaI, the reduced mass is

relatively heavy and the spread of the wave packet, a quantum mechanical feature, is only expected at longer times¹¹. More theoretical and experimental work is in progress.

The results reported here show that the motion of the nuclear wave packet during the breaking of the NaI bond can be probed with an experimental resolution of ~ 0.5 Å of internuclear separation. The position of the probing 'window' can be varied to map the trajectories $R(t)$. The packet is observed in real time for both contraction and extension of the NaI chemical bond as it passes through the window in both directions. The range of internuclear distances probed—up to 6.5 Å on the covalent and 12 Å on the ionic curves—shows the ability of the method to yield direct information on the potential-energy curves at large internuclear separations. This development promises a wide range of applications to other reactions, including those characterized by multidimensional potential-energy surfaces. □

Received 12 October; accepted 22 October 1990.

1. Zewail, A. H. *Science* **242**, 1645–1653 (1988).
2. Khundkar, L. R. & Zewail, A. H. *Rev. Phys. Chem.* **41**, 15–40 (1990).
3. Rose, T. S., Rosker, M. J. & Zewail, A. H. *J. chem. Phys.* **88**, 6672 (1988); **91**, 7415–7436 (1989).
4. Rosker, M. J., Rose, T. S. & Zewail, A. H. *Chem. Phys. Lett.* **146**, 175 (1988).
5. Atkins, P. W. *Physical Chemistry* 4th edn (Oxford University Press, New York, 1990).
6. Lee, S.-Y., Pollard, W. T. & Mathies, R. A. *J. chem. Phys.* **90**, 6146–6150 (1989).
7. Engel, V., Metiu, H., Almeida, R., Marcus, R. A. & Zewail, A. H. *Chem. Phys. Lett.* **152**, 1 (1988).
8. Bernstein, R. B. & Zewail, A. H. *J. chem. Phys.* **90**, 829–842 (1989).
9. Rosker, M. J., Dantus, M. & Zewail, A. H. *Science* **241**, 1200–1202 (1988).
10. Bower, R. D. et al. *J. chem. Phys.* **89**, 4478–4489 (1988).
11. Cong, P., Mokhtari, A. & Zewail, A. H. *Chem. Phys. Lett.* **172**, 109 (1990).
12. Engel, V. & Metiu, H. *J. chem. Phys.* **91**, 1596–1602 (1989).

ACKNOWLEDGEMENTS. This work was supported by the US Air Force.

Large variations in potential vorticity at small spatial scales in the upper ocean

R. T. Pollard* & L. Regier†

* Institute of Oceanographic Sciences, Deacon Laboratory, Brook Road, Wormley, Godalming, Surrey GU8 5UB, UK

† Scripps Institution of Oceanography, A-030, La Jolla, California 92093, USA

THE depth of the ocean's surface layer and stratification of the seasonal thermocline have long been thought to be controlled by heat exchange with the atmosphere and wind-induced mixing. Here we present observations which show that distortion by oceanic eddies also exerts an important influence on the structure of the upper ocean. Eddies induce horizontal shears, which are alternately cyclonic and anticyclonic on scales of 10–20 km. Conservation of potential vorticity then results in vertical motions which modify the stratification of the seasonal thermocline and can change its depth by 100 m or more in a few days. By subducting surface waters and transporting phytoplankton across the seasonal thermocline in a time comparable to that required to double the population, these motions are important to both the physics and the biology of the upper ocean.

One of us¹ has described the density structure of an oceanic front using data from an eight-leg survey in February 1986² obtained with a towed conductivity–temperature–depth measurement package (SeaSoar). The separation of isopycnals was very variable and there was a tendency for isopycnals in the thermocline beneath the surface outcrop of the front to deepen, suggestive of downward vertical velocities. We now have absolute velocity profiles from an acoustic Doppler current profiler (ADCP)^{3,4}, which can resolve similar scales to the SeaSoar (a few metres vertically, a few kilometres horizontally).

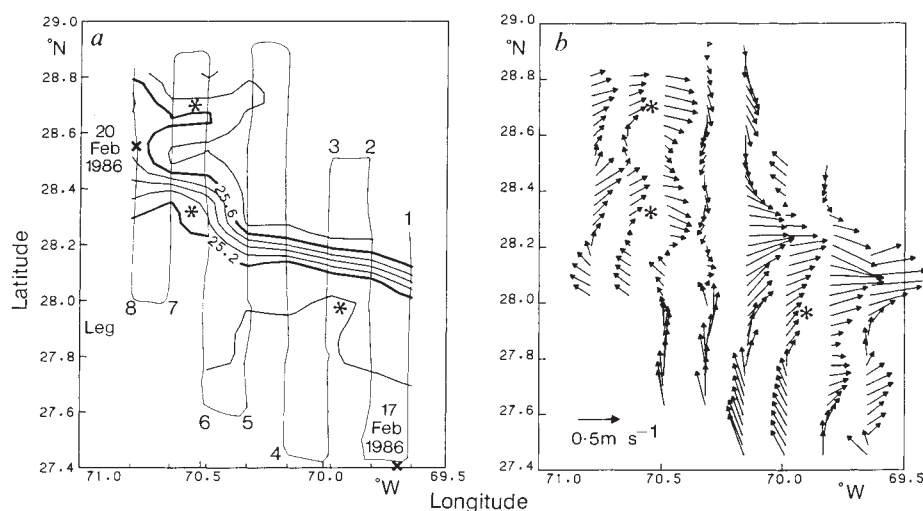


FIG. 1 Maps of *a*, density and *b*, velocity at a depth of 50 m, obtained from eight north-south sections across a front. The ship track, run by RV *Oceanus* during FASINEX², is shown in *a*. The velocity vectors, plotted every 4 km, are 10-km along-track averages. The three anticyclonic eddies in the density field (S-shaped features marked by asterisks) match anticyclonic shear in the velocity field. The shear $\partial u/\partial y$ alternates between being cyclonic and anticyclonic every 10 to 30 km along each north-south section.

The SeaSoar density and ADCP velocity measurements form a powerful combination to explore the small-scale dynamics of the upper ocean (Fig. 1). Where the SeaSoar shows small-scale eddies (20–40 km diameter) in the density field (Fig. 1*a*), the independent ADCP data show that the eddies produce strong horizontal shears ($3 \times 10^{-5} \text{ s}^{-1}$) in the velocity field (Fig. 1*b*). The alternating cyclonic and anticyclonic shears are apparent in all eight north-south sections, with the change from maximum to minimum velocity occurring in typically 20 km.

The key to understanding this structure is potential vorticity, which is the fluid equivalent of angular momentum. Just as an ice-skater speeds up her rate of spin by drawing in her arms, a column of water must become thinner and taller to increase its rate of rotation. Thus, conservation of potential vorticity leads to vertical velocities which transport properties across the surface layer. Conservation of potential vorticity⁵ holds from global⁶ and gyre⁷ scales down to scales of the order of 10 km (the internal Rossby radius of deformation in the ocean). Both the SeaSoar and ADCP resolve scales of the order of 4 km (ref. 1). The importance of potential vorticity Q to upper ocean dynamics on gyre scales has been discussed by Woods^{8,9}. Fischer *et al.*¹⁰ have recently mapped mesoscale variations of Q , and Bleck *et al.*¹¹ have modelled the vertical velocities that must arise at an ocean front.

Here we show sections across a front (Fig. 2) and maps (Fig. 3) to examine the contributions of stratification and relative vorticity to the potential vorticity field. Isopycnals (surfaces of constant density) commonly slope towards the surface, crossing from weak deep stratification through a strongly stratified thermocline into a weakly stratified surface layer. This results in large variations of Q along isopycnals, ultimately the product of air-sea interactions. These gradients are enhanced by eddies that strain the flow into alternating bands of high and low values of Q , like cream stirred in a coffee cup. Figure 1 shows shallow (100 m depth) 50-km eddies breaking across an ocean front. As they do so, they transport properties along isopycnals with slopes of 5×10^{-3} , so that there is a significant vertical component of the flow, necessary to conserve Q . Such structure is ubiquitous and will drive vertical motion important in the development of the seasonal thermocline and in the transport of properties across it.

The potential vorticity may be defined

$$Q = \frac{(f + \zeta) \Delta \rho}{\Delta p \rho}$$

where ζ is the relative vorticity and $\Delta \rho$ is the separation (thickness) in metres of two isopycnals $\rho \pm \Delta \rho/2$. This equation expresses the balance between the rate of rotation of the water

column, $f + \zeta$, and its height Δp . The rate of rotation must include both the Earth's vorticity f and the local relative vorticity ζ , which is usually much smaller than f . Contours of Q , Δp and ζ/f in two vertical sections are shown in Fig. 2, and isopycnal maps of Q , salinity S and ζ/f for $\rho = 25.45 \text{ kg m}^{-3}$ are shown in Fig. 3. The relative vorticity has been calculated from geostrophic velocity fields derived from the SeaSoar densities combined with a smoothed non-divergent velocity field at 150 m derived from the ADCP data.

Figure 2 shows how the alternating relative vorticity and varying thickness contribute to the potential vorticity. It is particularly clear in Fig. 2*a* how the hatched isopycnals are

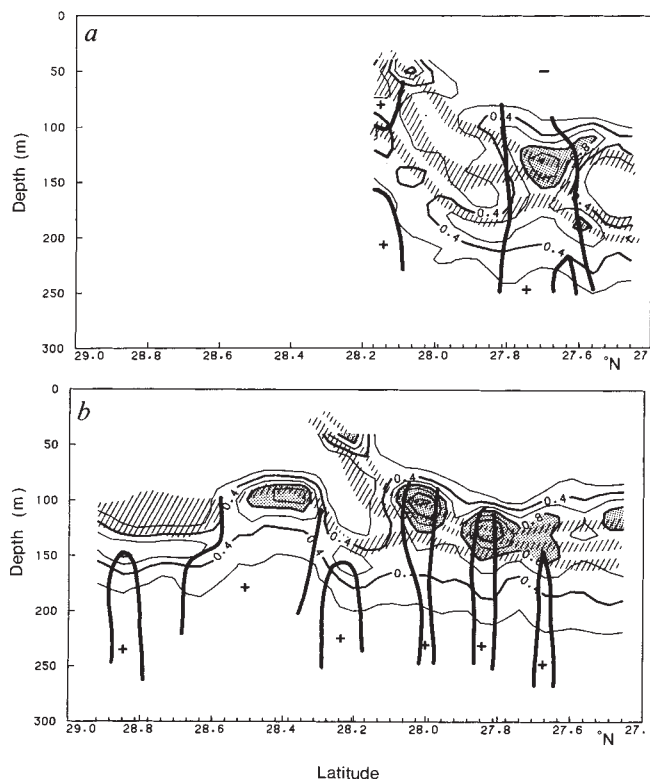


FIG. 2 Contours of Q at intervals of 0.2 PVU (here 1 PVU (potential vorticity unit) is $10^{-9} \text{ m}^{-1} \text{ s}^{-1}$) for *a*, leg 1 and *b*, leg 4. Hatched areas show the thickness of isopycnal bands centred at 25.45, 25.65 and 25.85 kg m^{-3} , and 0.1 kg m^{-3} wide. Areas of $Q > 0.8$ PVU are shaded, and areas of cyclonic relative vorticity ($\zeta > 0$) are marked (+) and bounded by thick lines.

squashed together within the broad band of cyclonic vorticity. This shows, from the equation for Q , that both diminished thickness and enhanced vorticity are contributing to the potential vorticity maximum between 120–150 m in the seasonal thermocline. Thus stratification and relative vorticity are positively correlated to enhance variations in Q , which is the typical signature of advected anomalies of potential vorticity^{5,12}.

The correlation changes with depth. In the 25.8–25.9 kg m^{-3} isopycnal band, the correlation of stratification and relative vorticity has become negative. Thick regions correspond to regions of cyclonic vorticity (Fig. 2), so that the variation of Q along isopycnals is weak. In the 26.2–26.3 kg m^{-3} isopycnal band (not shown), lying at depths between 200 and 250 m, the total range of Q is only 0.1 PVU, the variations in relative vorticity being almost entirely cancelled by opposing variations in stratification. These observations confirm that the strong variations in stratification are controlled, first by air-sea interaction setting up the change in stratification along isopycnals just below the base of the mixed layer and outcropping into the mixed layer, and second by conservation of Q as eddies strain the Q field to smaller and smaller scales.

To give confidence that we have resolved potential vorticity satisfactorily, we have mapped both salinity (Fig. 3a) and potential vorticity (Fig. 3b). Both Q and S are conservative tracers on isopycnals in the absence of mixing, and the strong correla-

tion between them is apparent. Both show east-west orientated bands of alternate high and low values, with maxima and minima as close as 10 km apart in the north-south direction. Q shows a tenfold variation, from 0.15–1.5 PVU between 28.0 and 28.1° N. This tenfold variation is very large, as large as the range found on the 25.5 kg m^{-3} isopycnal for the whole North Atlantic¹³. In this example, most of the change in Q is provided by a change in relative vorticity from 0.2f to -0.74f (Fig. 3c). There is thus a fivefold change in absolute vorticity ($f + \zeta$), from 1.2 to 0.26, enhanced by a change of a factor of two in thickness.

The way in which this 10-km-scale structure arises is as follows. Envisage an initial state in which there is a front lying east-west. Wind mixing enhances the stratification just below the base of the mixed layer, so that, in that depth range only, there are isopycnals along which there is a strong change in Q , from large values in the thermocline (Fig. 2, shaded areas) to small values where isopycnals outcrop into the weakly stratified surface layer. Eddies now distort the front, initially into S shapes (Fig. 1) eventually straining the flow into alternating bands of high Q and low Q , whose widths become smaller and smaller until they are limited by the Rossby radius of deformation. Banded structure in mesoscale (100 km) eddies and indications of many small-scale (20 km) eddies are common features in satellite images, so we believe our data are representative of ubiquitous oceanic motions.

As the scale of the alternating bands decreases, any shear in the water column will carry water parcels across a strong horizontal gradient of relative vorticity. The relative vorticity of the water parcel must adjust to match that of its surroundings (note how the contours of constant relative vorticity tend to be vertical in Fig. 2). As it does so, its thickness must also change, to conserve its potential vorticity. This leads to vertical velocities, which we estimate can be as large as tens of metres per day. For example, tongues of high Q can be seen on the 25.45 kg m^{-3} surface where it outcrops (Fig. 3a), rising towards the surface between legs 1 and 2 and legs 5 and 6. We estimate that the tongues rise 50 m as they move 20 km along the front at 0.2 m s^{-1} , giving a vertical velocity of $50 \times 10^{-5} \text{ m s}^{-1}$, or 50 m d^{-1} . Such large velocities are of considerable importance to the evolution of the thermocline, its acoustic properties and its biological productivity⁹. Mixed-layer depth can double locally from 100 m to 200 m in a few days. Phytoplankton can be carried out of the euphotic zone in a time similar to the time required to double the population. The small-scale patchiness of phytoplankton in the surface layer is directly related to the physical processes we have described.

Eventually, prediction of the physical and biological evolution of the upper ocean will have to rely on satellite imagery for the initial values required by computer models. An important objective of this work is therefore to interpret the detailed structure seen on satellite images in terms of the sub-surface physical processes that create it. □

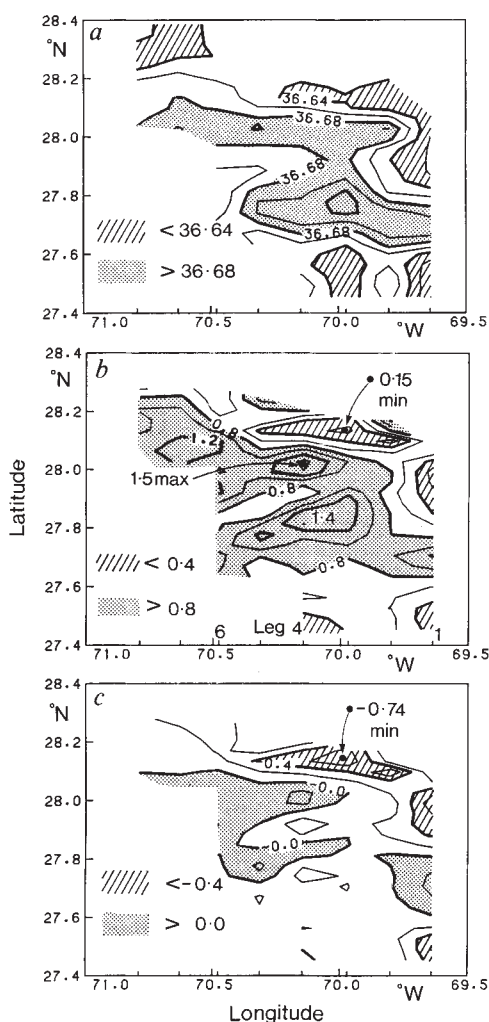


FIG. 3 Maps of a, potential vorticity Q , b, salinity S and c, isopycnal relative vorticity ζ/f on the surface $\rho = 25.45 \text{ kg m}^{-3}$. This surface is the upper hatched band in Fig. 2.

Received 28 December 1989; accepted 9 October 1990.

- Pollard, R. T. *Nature* **323**, 433–435 (1986).
- Stage, S. A. & Weller, R. A. *Bull. Am. met. Soc.* **66**, 1511–1520 (1985); **67**, 16–20 (1986).
- Regier, L. *J. phys. Oceanogr.* **12**, 880–886 (1982).
- Pollard, R. T. & Read, J. F. *J. atmos. ocean. Technol.* **6**, 859–865 (1989).
- Hoskins, B. J., McIntyre, M. E. & Robertson, A. W. *Q. J. R. met. Soc.* **111**, 877–946 (1985).
- McIntyre, M. E. & Palmer, T. N. *Nature* **305**, 593–600 (1983).
- Rhines, P. B. A. *Rev. Fluid Mech.* **18**, 433–397 (1986).
- Woods, J. D. in *Coupled Ocean—Atmosphere Models* (ed. Nihoul, J. C. J.) 543–590 (Elsevier, Amsterdam, 1985).
- Woods, J. D. in *Toward a Theory of Biological-Physical Interactions in the World Ocean* (ed. Rothschild, B. J.) 7–38 (Kluwer, Dordrecht, 1988).
- Fischer, J., Leach, H. & Woods, J. D. *J. phys. Oceanogr.* **19**, 519–531 (1989).
- Bleck, R., Onken, R. & Woods, J. D. *Q. J. R. met. Soc.* **114**, 347–371 (1988).
- McIntyre, M. E. in *The Nature and Prediction of Extratropical Weather Systems*, 237–259 (European Centre for Medium-range Weather Forecasting, Reading, 1988).
- Stammer, D. & Woods, J. D. *Isopycnal Potential Vorticity Atlas of the North Atlantic Ocean*, *Berichte aus dem Inst. für Meereskunde*, Kiel **165** (1987).

ACKNOWLEDGEMENTS. This work was supported by the US Office of Naval Research and the National Environment Research Council.

Radiocarbon evidence of fossil-carbon cycling in sediments of a nearshore hydrocarbon seep

James E. Bauer*, Robert B. Spies†, John S. Vogel‡§, D. Erle Nelson‡ & John R. Southon‡§

* Marine Research Division, Scripps Institution of Oceanography, University of California, La Jolla, California 92093, USA

† Environmental Sciences Division, University of California, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

‡ Department of Archaeology, Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6

§ Present Address: Center for Accelerator Mass Spectrometry, University of California, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

WIDESPREAD seepage of petroleum and natural gas in the southern California continental borderland¹⁻⁴ provides an opportunity to study the long-term fate and biogeochemical effects of hydrocarbons in nearshore sedimentary environments. The hydrocarbons that enrich seep sediments have been hypothesized to serve as a carbon and energy source for sediment metabolism and infaunal populations⁵⁻⁹. Here we present ¹⁴C natural abundances in sediment total organic carbon (TOC), pore-water dissolved inorganic carbon (DIC) and infauna from in and around a hydrocarbon seep off southern California which were measured to help test this hypothesis. Concentrations of ¹⁴C in each pool reflect the admixture of fossil (¹⁴C-depleted) seep-derived carbon with carbon from the euphotic zone. The ¹⁴C depletion in TOC and DIC increased with proximity to the seepage zone and with sediment depth; ¹⁴C abundances differed between meiofauna and macrofauna, suggesting that the two groups incorporate fossil carbon in different ways. The results indicate that fossil carbon can indeed comprise a major component of these carbon pools in nearshore seep sediments.

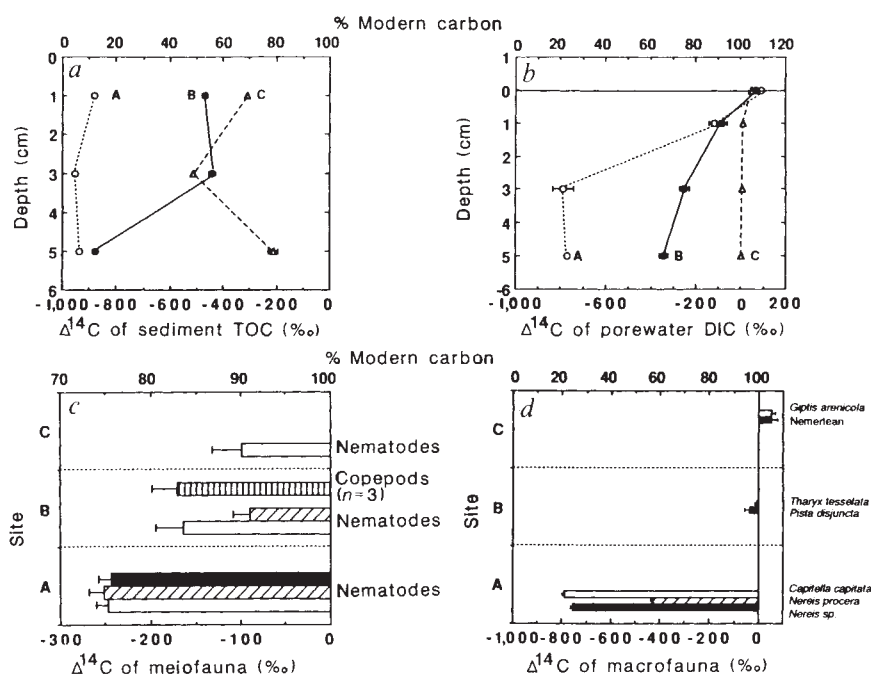
The Isla Vista hydrocarbon seep (34°19' N, 119°50' W), located offshore from Santa Barbara, differs from the deeper, well studied seeps of the Gulf of Mexico^{10,11} in its shallow depth

(16 m) and nearshore location (~0.5 km) and in the absence of vent-type fauna. The area is characterized by fine, well sorted Quaternary sands which are directly underlain by highly fractured Miocene Monterey shale, and the major seep trends in the region lie along active strike-slip and thrust faults^{2,4,12}. Total rates of petroleum seepage for the Coal Oil Point region, of which the Isla Vista seep is a discrete seep, range from 15 to 400 barrels per day⁴. Sediment hydrocarbons at the Isla Vista seep consist of abundant natural gases (C₁-C₅) and higher molecular weight aliphatic and aromatic compounds^{13,14} (J.E.B., manuscript in preparation).

We obtained samples of TOC, DIC, meiofauna and macrofauna for ¹⁴C analysis from cores collected at three sites: A, at the centre of the Isla Vista seep where extensive petroleum and hydrocarbon gas seepage occur; B, 20 m to the north of A on the periphery of the seep where there is little oil and gas evolved but large numbers of weathered tar particles; C, 1 km to the east of A and B, which served as a control site. In nearshore environments, ¹⁴C is a more useful carbon tracer for biogeochemical and trophic-dynamic studies than is ¹³C because its sources are distinct and because it has a greater natural dynamic range¹⁵. We measured the ¹⁴C abundances of 58 samples, some as small as 25 µg, by accelerator mass spectrometry to precisions of 1-10% depending on their size¹⁶. Although the qualitative nature of carbon and energy flow through Isla Vista seep sediments was inferred from previous studies^{8,9} of community metabolism, quantitative estimates were not possible. Here we use natural abundances of ¹⁴C to provide estimates of the extent of fossil-carbon cycling in the major carbon pools of seep and non-seep sediments.

The quantity and ¹⁴C content of TOC at the three sites provides a basis for estimating the type of sediment organic carbon available for subsequent sediment metabolism and faunal incorporation (Fig. 1a; Table 1). The depth-averaged (0-6 cm) ¹⁴C of TOC at site C (mean = 65.6 ± 15.4% modern, modern ¹⁴C is defined as the 1950 AD atmospheric ¹⁴C concentration) was comparable to other California borderland surficial sediments, including those from the Santa Barbara (0-19 cm; 72.1% modern) and San Nicolas (0-5 cm; 70.3% modern) basins¹, but

FIG. 1 Natural ¹⁴C abundances, expressed as percent modern carbon and $\Delta^{14}\text{C}$, in a, sediment TOC, b, pore-water DIC, c, meiofauna and d, macrofauna in sediments at sites A, B and C near the Isla Vista hydrocarbon seep. Values for TOC and DIC represent the mid-point of the depth interval sampled. Organisms were collected from the 0-2 cm depth interval only. Bars indicate a measurement error of 1 standard deviation for individual samples (TOC, DIC and macrofauna) or composite samples (meiofauna). Sediments were dried, treated with dilute HCl to remove carbonates and stored dry and in the dark. Pore waters were collected by filter-centrifuging core sections through combusted GF/F glass-fibre filters and stored under N₂ at 5 °C in the dark using HgCl₂ as a preservative. Organisms were separated from surface sediments (0-2 cm) by elutriation and hand-sorting into clean 0.2-µm-filtered sea water where they were allowed to clear their guts for 24 h. Animals were then transferred to distilled water and dried at 60 °C for 48 h. Sediment and faunal samples were subsequently combusted at 900 °C, and pore waters were acidified with concentrated phosphoric acid. The evolved CO₂ was reduced under an atmosphere of H₂ to graphite²⁶, which was subsequently analysed using the Simon Fraser University accelerator mass spectrometry facility at McMaster University²⁷. Measurements of ¹³C content, required for ¹⁴C normalization, were not made directly on the samples owing to the small amounts of material available. Instead we used published measurements^{1,20-22,28,29} for ¹³C content representative of each sample type. The ¹⁴C determinations were not affected by uncertainties



in ¹³C content within the total precision of the radiocarbon measurements. Radiocarbon contents and apparent ages of the samples are expressed in conventional notation³⁰. $\Delta^{14}\text{C}$ is the per mil deviation from the 'standard' nineteenth century wood.

TABLE 1 Profiles of Isla Vista seep sediments

Site	Depth (cm)	TOC (%)	DIC (mM)	Bacteria ($\times 10^8 \text{ cm}^{-3}$)	^{14}C Activity		TOC oil fraction (%)	Respired DIC (%)
					TOC (% modern)	DIC (% modern)		
A	0	—	2.2 $\pm$ 0	—	—	109.3 $\pm$ 0.2	—	—
	0–2	1.8	7.4 $\pm$ 1.7	32.0 $\pm$ 3.4	12.1 $\pm$ 0.3	88.5 $\pm$ 2.5	84	14
	2–4	1.9	7.6 $\pm$ 2.2	37.0 $\pm$ 0.6	4.6 $\pm$ 0.3	21.4 $\pm$ 4.7	92	81
	4–6	1.8	9.4 $\pm$ 0.9	51.0 $\pm$ 13.0	6.2 $\pm$ 0.3	23.0 $\pm$ 0.8	90	79
B	0	—	2.2 $\pm$ 0	—	—	107	—	—
	0–2	0.3	3.7 $\pm$ 0.9	20.0 $\pm$ 4.1	53.0 $\pm$ 1.0	91.8 $\pm$ 2.4	24	25
	2–4	1.6	2.9 $\pm$ 0.7	13.0 $\pm$ 3.0	56.0 $\pm$ 1.3	75.0 $\pm$ 2.1	40	67
	4–6	2.8	4.3 $\pm$ 0	8.6 $\pm$ 2.0	12.4 $\pm$ 0.4	65.9 $\pm$ 2.0	90	43
C	0	—	2.2 $\pm$ 0	—	—	104.9 $\pm$ 1.9	—	—
	0–2	0.2	2.2 $\pm$ 0	9.0 $\pm$ 1.8	69.1 $\pm$ 0.8	101.0 $\pm$ 4.1	—	11
	2–4	0.2	2.2 $\pm$ 0.3	8.0 $\pm$ 1.4	48.8 $\pm$ 0.8	100.7 $\pm$ 4.4	—	7
	4–6	0.2	2.2 $\pm$ 0.1	6.6 $\pm$ 2.0	78.9 $\pm$ 1.7	100.4 $\pm$ 2.6	—	14

TOC and DIC were determined using standard methods²⁴; bacteria were enumerated by epifluorescence microscopy²⁵. The fraction of TOC derived from petroleum carbon (0% modern C) for each depth at sites A and B was estimated using a mass-balance relationship which assumed that linear mixing of average site-C-type sediment (mean TOC=0.2%, mean ^{14}C =65.6 $\pm$ 15.4% modern) TOC with petroleum carbon results in TOC having the measured activities at sites A and B. The maximum potential fraction of DIC derived from respired organic carbon at each depth was also calculated by mass balance assuming that linear mixing of organic-carbon-derived DIC with sediment surface water (0 cm) DIC takes place to give the observed pore-water DIC concentration and ^{14}C activity at sites A, B and C. Values for TOC at all sites and for DIC at site C were determined previously⁹. TOC, DIC and bacteria are means of replicate cores with an uncertainty of 1 standard deviation. DIC ^{14}C activity at 0 cm at site B was not measured but was assumed to be the average of the activities measured at 0 cm at sites A and C.

did not show a regular ^{14}C pattern with depth. Assuming that the average sediment at site C typifies non-seep sediments in the region, the ^{14}C concentrations of sediments at sites A and B may be used to estimate the amount of petroleum carbon in site A and B sediments (Table 1). The TOC at site A contained 84–92% seep-derived, non-volatile carbon and was depleted in ^{14}C at even the shallowest depths. The TOC from site B resembled normal borderland (site C) sediment near the surface but was similar to the hydrocarbon-rich sediment of site A at 4–6 cm.

The abundance of ^{14}C in DIC (Fig. 1b) allows us to estimate the maximum potential remineralization of sediment organic carbon. We assumed that DIC derived from the complete oxidation of 'average' bulk TOC at each depth mixes linearly with DIC from the sediment/seawater interface (Table 1). At site C, advective and diffusive mixing of pore-water DIC with DIC from overlying sea water limited the steady-state contribution of DIC from TOC remineralization to no more than 14%. If we assume that DIC is produced locally in seep (sites A and B) sediments, up to ~80% of the steady-state DIC in site A pore waters is potentially derived from the remineralization (presumably mediated by high rates of sulphate reduction at site A⁹) of TOC and pore-water dissolved organic compounds (J.E.B., manuscript in preparation). In contrast, up to ~70% of DIC from site B may be derived from organic carbon respiration, but this is under suboxic rather than sulphate reducing conditions⁹. Another potential mechanism contributing to elevated fossil DIC in pore waters is its transport from the deep sub-surface reservoir to shallow sediments in formation waters¹⁷. Although our data show significant concentrations of fossil DIC near the seep centre, they do not distinguish between potential sources of DIC as all ultimately arise from fossil organic carbon. The high DIC concentrations, bacterial abundances (Table 1) and rates of metabolism^{8,9} at site A suggest, however, that at least a portion of the fossil DIC in these sediments is produced by recent remineralization of organic carbon.

Nematodes comprised the bulk of meiofauna sampled, and harpacticoid copepods were present in sufficient quantities for ^{14}C measurement at site B only (Fig. 1c). Meiofauna (63–500 μm equivalent spherical diameter, 0.8–1.0 μg per organism) feed by selecting material from the surfaces of sediment grains or in sediment interstices.¹⁸ The ^{14}C contents of meiofauna at site C indicate that their primary source of nutritive carbon was contemporary. Without specifying trophic or assimilatory pathways, we deduced the net incorporation of sedimentary, and hence

fossil, organic carbon at the seep (sites A and B). To estimate the fraction of meiofauna carbon derived from sedimentary TOC we assumed that organisms incorporated carbon from a mixture of sediment TOC and euphotic organic carbon, which was similar in ^{14}C content to the surface DIC¹⁹ (Table 2). The meiofauna from all three sites obtained 60–65% of their carbon from other than average bulk TOC, whether the sediment carbon was 82% fossil (at site A) or was typical borderland sediment (at site C). Stable-carbon isotope evidence indicates that meiofauna can select specific subfractions of sediment organic matter^{20–22}, which, in the seep environment, may include more assimilable and nitrogen-rich organic matter²³ than hydrocarbons. Alternatively, meiofauna may not feed in the most hydrocarbon-rich microenvironments to avoid effects of toxicity. Either scenario suggests that certain meiofauna regulate their incorporation of carbon in hydrocarbon-rich environments.

The ^{14}C abundances of deposit-feeding macrofauna (primarily polychaetes) were more variable than meiofauna (Fig. 1d). The fraction of their carbon derived from fossil carbon (Table 2) equalled the fraction of oil in surface sediment (0–2 cm) TOC (Table 1), within reasonable uncertainty. Site C macrofauna were not depleted in ^{14}C at all with respect to hypothetical euphotic organic matter suggesting, along with the ^{14}C DIC data for site C (Table 1, Fig. 1b), that this TOC is refractory and of low nutritive value, accounting for its mean 3,550-yr equivalent age. Of the organisms analysed, the greatest amounts of fossil carbon were incorporated by site A polychaetes

TABLE 2 ^{14}C in fauna and surface sediment

	A	B	C
Surface sediment ^{14}C (% modern C)	12.1 $\pm$ 0.3	53.0 $\pm$ 1.0	69.1 $\pm$ 0.8
Meiofauna ^{14}C (% modern C)	75.3 $\pm$ 0.7	84.5 $\pm$ 4.2	90.1 $\pm$ 3.3
Carbon from sediment TOC (%)	35	37	41
Macrofauna ^{14}C (% modern C)	34.7 $\pm$ 14.0	98.2 $\pm$ 2.1	105.2 $\pm$ 1.6
Carbon from sediment TOC (%)	77	16	0

The percentage of carbon derived from surface-sediment TOC in each faunal group was estimated by assuming that organisms use carbon from a mixing line with the carbon of surface sediment TOC and sediment-supplied organic carbon (assumed to be similar in radiocarbon signature to sediment-surface DIC¹⁹) as endmembers.

(Fig. 1d). The data show that macrofauna, unlike meiofauna, do not discriminate among sources of organic matter with different fossil-carbon contents and sequester their body carbon in direct proportion to sediment TOC, regardless of the specific quality of the organic matter (for example, C/N is ~ 20 at site A⁹).

The large number of documented hydrocarbon seeps in the southern California continental borderland²⁻⁴ indicates that the extensive biogeochemical cycling of hydrocarbon-derived carbon described here in Isla Vista seep sediments could be a widespread regional, if localized, phenomenon. Rates of important geochemical processes including community oxygen meta-

bolism and sulphate reduction are significantly enhanced in seep sediments⁹. Although sediments at the Isla Vista seep have similar types of organisms as non-seep coastal sediments, the numbers of organisms and the distribution of individual taxa are significantly altered⁵⁻⁷. We suggest that these processes and distributions result, at least in part, from extensive incorporation of fossil carbon into seep sediments and its subsequent use as a carbon and energy source by sedimentary assemblages. Careful extrapolation of the information derived from studies of natural hydrocarbon seepage may also provide information on the fate and effects of chronic hydrocarbon inputs to other nearshore sedimentary environments. □

Received 30 May; accepted 27 September 1990.

- Emery, K. *The Sea Off Southern California* (Wiley, New York, 1960).
- Wilkinson, E. *Calif. Div. Oil Gas Publ. No. TRO8*, 1-11 (1972).
- Reed, W. & Kaplan, I. *Geochem. Explor.* **7**, 255-293 (1977).
- Fisher, P. *Calif. State Lands Comm. Rep.* 1-62 (1978).
- Davis, P. & Spies, R. *Mar. Biol.* **59**, 31-41 (1980).
- Montagna, P. *et al. J. mar. Res.* **45**, 761-789 (1987).
- Montagna, P., Bauer, J., Hardin, D. & Spies, R. *J. mar. Res.* **47**, 657-680 (1989).
- Montagna, P. *et al. Mar. Ecol. Prog. Ser.* **34**, 31-40 (1986).
- Bauer, J. *et al. Limnol. Oceanogr.* **33**, 1493-1513 (1988).
- Brooks, J. *et al. Science* **238**, 1138-1142 (1987).
- Paull, C. *et al. Nature* **342**, 166-168 (1989).
- Wilson, R. *et al. Science* **184**, 857-865 (1974).
- Spies, R., Davis, P. & Steurmer, D. in *Marine Environmental Pollution Vol. 1* (ed. Geyer, R.) 229-263 (Elsevier, New York, 1980).
- Steurmer, D. *et al. Mar. Chem.* **11**, 413-426 (1982).
- Degens, E. in *Organic Geochemistry* (eds Eglinton, G. & Murphy, M.) 304-329 (Springer-Verlag, New York, 1970).
- Vogel, J., Briskin, M., Southon, J. & Nelson, E. *Radiocarbon* (in the press).

- Collins, A. *Geochemistry of Oilfield Waters* (Elsevier, New York, 1975).
- Reise, K. *Tidal Flat Ecology* (Springer, Berlin, 1985).
- Eppeley, R., Renger, E. & Betzer, P. *Deep Sea Res.* **30**, 311-323 (1983).
- Spies, R. & Des Marais, D. *Mar. Biol.* **73**, 67-71 (1983).
- Schwinghammer, P., Francis, F. & Gordon, D. *Can. J. Fish. Aquat. Sci.* **40** (Suppl. 1), 262-272 (1983).
- Couch, C. *Estuar. Coast. Shelf Sci.* **28**, 433-441 (1989).
- Findlay, S. & Tenore, K. *Science* **218**, 371-373 (1982).
- Strickland, J. & Parsons, T. *Bull. Fish. Res. Bd. Can.* **167** (1972).
- Daley, R. & Hobbie, J. *Limnol. Oceanogr.* **20**, 875-882 (1975).
- Vogel, J., Nelson, E. & Southon, J. *Radiocarbon* **29**, 323-333 (1987).
- Nelson, E., Vogel, J., Southon, J. & Brown, T. *Radiocarbon* **28**, 215-222 (1986).
- Emery, K. & Bray, E. *Bull. Am. Ass. Petrol. Geol.* **46**, 1839-1856 (1962).
- Nissenbaum, A., Presley, B. & Kaplan, I. *Geochim. cosmochim. Acta* **36**, 1007-1027 (1972).
- Stuiver, M. & Polach, H. *Radiocarbon* **19**, 355-363 (1977).

ACKNOWLEDGEMENTS. We thank S. Anderson and J. McCullagh for technical assistance and P. M. Williams, G. Rau and S. Trumbore for comments on earlier versions of this manuscript. Financial support was provided by the Office of Institutional Research and Development at Lawrence Livermore National Laboratory and the Natural Science and Engineering Research Council of Canada.

Evidence from chronosequence studies for a low carbon-storage potential of soils

William H. Schlesinger

Departments of Botany and Geology, Duke University, Durham, North Carolina 27706, USA

OVER most of the Earth's land surface, the amount of carbon stored in soil organic matter exceeds by a factor of two or three the amount stored in living vegetation. This pool of soil carbon is large (1.5×10^{18} g)^{1,2} and plays a dynamic part in the geochemical carbon cycle. Prentice and Fung³ have suggested that terrestrial vegetation and soils would act as a large sink for atmospheric carbon dioxide if its concentration were twice the present level. Here I use data from chronosequence studies to show that the production of refractory humus substances in soils sequesters only $\sim 0.4 \times 10^{15}$ g C yr⁻¹ from the atmosphere, accounting for just 0.7% of terrestrial net primary production. Moreover, agricultural practices tend, on balance, to cause a release of soil carbon to the atmosphere^{4,5}. Thus if the terrestrial biosphere is indeed to act as a carbon sink under future elevated levels of carbon dioxide, this would be more likely to be the result of changes in the distribution and biomass of terrestrial vegetation than of changes in the accumulation of soil organic matter.

The carbon stored in soil organic matter represents the long-term net balance of photosynthesis and total respiration in terrestrial ecosystems. Radiocarbon dates of humic materials in the lower soil profile are frequently $>1,000$ years⁶. New land surfaces produced by disturbances such as volcanic eruptions or glacial retreat yield youthful soils with little or no organic matter. In these soils, the long-term average rate of accumulation of soil organic matter is easily calculated from measurements of bulk density and the percentage of organic carbon at each depth in the soil profile of depth N . Soil organic carbon (g C m^{-2}) = $\sum_{D=0}^N 10,000 D \rho(D) P(D)$ where $\rho(D)$ and $P(D)$ are the bulk density in g cm^{-3} and the percentage of organic carbon,

respectively, in depth interval D . The long-term average rate of accumulation is obtained by dividing the carbon content of the profile by the estimated or measured age. Interval-specific accumulation rates can be calculated from the difference between two soil profiles in the same chronosequence, divided by the difference in their ages.

During soil development, organic matter often shows an initial period of rapid increase for up to 3,000 years, followed by a lower rate of accumulation that may continue for millennia⁷. At any time, the long-term average rate of accumulation is an overestimate of the current rate of accumulation in most soils (Fig. 1). As most soils are at least 3,000 years old, however, the average long-term rate for 3,000- to 10,000-year-old soils provides an upper limit for estimates of soil carbon accumulation during the Holocene and in possible future climates.

In upland ecosystems, the long-term rate of carbon storage varies from $0.2 \text{ g C m}^{-2} \text{ yr}^{-1}$ in some polar deserts to $>10 \text{ g C m}^{-2} \text{ yr}^{-1}$ in some forests (Table 1). Considering the paucity of studies and the range of environments represented, the rates are surprisingly uniform. The long-term rate for 16 soils that are 3,000 to 10,000 years old, $2.4 \pm 0.70 \text{ g C m}^{-2} \text{ yr}^{-1}$ (1 s.e.) (Table 1), can be used to calculate the maximum potential for the present accretion of soil organic matter.

An annual production of $2.4 \text{ g C m}^{-2} \text{ yr}^{-1}$ of humic substances implies a total storage of $0.32 \times 10^{15} \text{ g C yr}^{-1}$ in upland ecosystems under natural vegetation ($133 \times 10^6 \text{ km}^2$; ref. 5). The actual net storage of organic carbon in upland soils is probably less than $0.32 \times 10^{15} \text{ g C yr}^{-1}$, because this value assumes that all upland soils are only 3,000-10,000 years old. Following a similar approach, Armentano and Menges⁸ found that before recent disturbance, soils in wetlands, bogs and peatlands ($3.5 \times 10^6 \text{ km}^2$) were a sink for up to $0.08 \times 10^{15} \text{ g C yr}^{-1}$. Their value may be too low, given recent higher estimates of global wetland area⁹. Nevertheless, assuming these estimates are reasonable, the total storage of soil organic matter ($0.40 \times 10^{15} \text{ g C yr}^{-1}$) is $\sim 0.7\%$ of recent estimates of net primary production on land ($60 \times 10^{15} \text{ g C yr}^{-1}$; ref. 10), but 20% of that storage occurs in wetland ecosystems which cover only 2% of the land surface. A similar approach shows that the net storage of carbon in desert-soil carbonate is also a minor flux in the global carbon cycle ($<0.023 \times 10^{15} \text{ g C yr}^{-1}$; ref. 11).

TABLE 1 Long-term rates of accumulation of organic carbon in Holocene-age soils

Ecosystem type	Vegetation in terminal state	Soil origin	Accumulation interval (yr)	Long-term rate of accumulation ($\text{g C m}^{-2} \text{ yr}^{-1}$)	Ref.
Tundra	Polar desert	Glacial retreat	8,000	0.2	32
	Polar desert	Glacial retreat	9,000	0.2	33
	Polar desert	Glacial retreat	2,600	2.4	34
	Sedge moss	Glacial retreat	1,000	2.4	33
	Sedge moss	Glacial retreat	9,000	1.1	33
	Sedge moss	Glacial retreat	8,700	5.7	34
Boreal forest	Spruce	Glacial retreat	3,500*	11.7	35
	Spruce-fir	Glacial retreat	5,435	0.8	16
	Spruce-fir	Glacial retreat	2,740	2.2	17
Temperate forest	Broadleaf evergreen	Volcanic ash	1,277	12.0	36
	Coniferous	Volcanic mudflow	1,200	10.0	27
	Deciduous	Alluvium	1,955	5.1	37
	Deciduous	Dunes	10,000	0.7	38
	<i>Podocarpus</i>	Dunes	10,000	2.1	28
	<i>Angophora</i>	Dunes	4,200	1.7	39
	<i>Eucalyptus</i>	Dunes	6,500	1.4	40
	<i>Eucalyptus</i>	Dunes	5,500	2.1	41
	Low forest	Glacial deposits	9,000	2.5	42
	<i>Metrosideros</i>	Volcanic ash	3,500	2.5	43
Tropical forest	Rain forest	Volcanic ash	8,620	2.3	44
Temperate grassland	<i>Chionochloa</i>	Glacial deposits	9,000	2.2	42
Temperate desert	Grassland	Alluvium	3,040	0.8	45

Unless actual data were included, calculated accumulation rates assume soil bulk density of 1.5 g cm^{-3} and that soil organic matter contains 50% organic carbon by mass. Rates do not include accumulations of undecomposed organic matter on the soil surface.

* Data corrected for 8 kg C m^{-2} in forest floor and a new estimate for soil age⁴⁶.

On a unit-area basis, the annual production of refractory humus substances in upland soils is greater than the deposition of carbon in deep-ocean sediments ($0.44 \text{ g C m}^{-2} \text{ yr}^{-1}$; ref. 12), but ocean sediments cover a greater area and are likely to accumulate and persist for longer periods of time. Organic matter does not accumulate in upland soils forever; eventually a steady-state profile content is achieved. In many forests the soil profile contains $\sim 10\text{--}20 \text{ kg C m}^{-2}$ (refs 1, 2), implying that $>6,000$ years are required to establish a stable distribution of carbon in the soil. At this point the production of humic compounds, in excess of decomposition, must equal the removal of these substances by erosion. Thus, estimates of the transport of organic carbon in rivers provide an upper limit for the production of refractory humic substances in the soil (some river transport is derived from fresh plant material that enters rivers directly). Recent estimates of the global transport of organic carbon in rivers ($0.4 \times 10^{15} \text{ g C yr}^{-1}$; ref. 13) are consistent with the estimate of an annual production of $0.40 \times 10^{15} \text{ g C yr}^{-1}$ of soil humus substances in steady-state conditions. Some of the organic carbon carried by rivers is degraded in the sea¹² and the remainder is deposited in sediments of estuaries and on the continental shelf¹⁴.

At the maximum extent of Pleistocene glaciation, $29.5 \times$

10^6 km^2 of the present land area was covered with ice¹⁵. This area now contains roughly $400 \times 10^{15} \text{ g C}$ or $\sim 25\%$ of the carbon contained in all soils of the world¹. The pool of carbon in deglaciated soils is consistent with accumulation rates of $<2.4 \text{ g C m}^{-2} \text{ yr}^{-1}$ for the last 10,000 years^{16,17}. The annual rate of storage in glaciated ecosystems ($0.04 \times 10^{15} \text{ g C yr}^{-1}$) is too small to have affected atmospheric CO_2 significantly during the Holocene³. It is difficult to see how this rate of storage might have changed enough during the past century to account for the $2\text{--}3.4 \times 10^{15} \text{ g C yr}^{-1}$ that is currently missing from the atmospheric inventory of CO_2 ¹⁸. Moreover, if the global climate becomes warmer, most soils are likely to become net sources of atmospheric CO_2 , because higher soil temperatures increase decomposition rates^{19–21}. Shifts in the distribution and extent of terrestrial vegetation will also produce changes in the storage of soil organic carbon. The areal extent of vegetation predicted by Emanuel *et al.*²² and the soil carbon contents of Post *et al.*² suggest a loss of $45.5 \times 10^{15} \text{ g C}$ from the soil carbon pool when the atmospheric CO_2 concentration is double the present value. In a similar analysis, Prentice and Fung³ suggest a slight increase in the soil carbon pool in the ecosystems of a warmer Earth.

Most soils show net losses of organic matter when they are converted to agricultural use⁵ but soil carbon contents increase

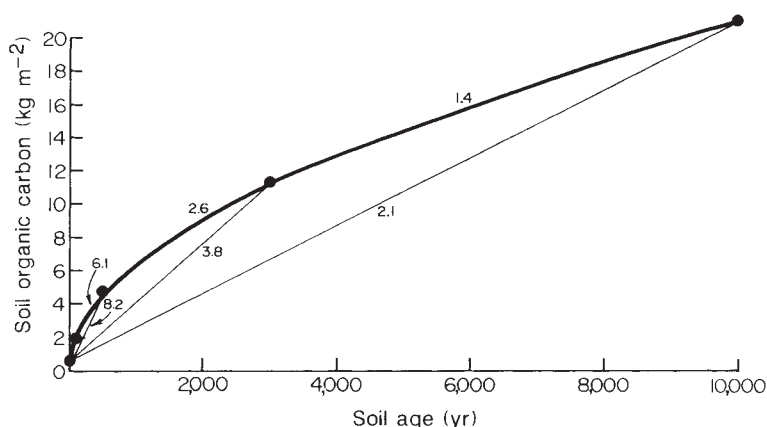


FIG. 1 The rate of carbon accumulation in New Zealand sand dunes²⁸. For any soil age, the calculated long-term rate of accumulation overestimates the actual rate of accumulation during the most recent interval.

when these lands are allowed to return to native vegetation^{23,24}. Soil organic matter may also be enhanced under some types of improved agricultural management²⁵. Between 1950 and 1959, $27 \times 10^4 \text{ km}^2$ of farmland was abandoned in the eastern United States²⁶. Using $30 \text{ g C m}^{-2} \text{ yr}^{-1}$ as an estimate of the rate of accumulation of soil organic matter in 40- to 50-year-old soils^{24,27,28} results in a storage of $0.008 \times 10^{15} \text{ g C yr}^{-1}$ in these soils—an insignificant flux in the global carbon cycle. Similarly, Delcourt and Harris²⁹ report that forests in the southeastern United States have served as a sink for $0.07 \times 10^{15} \text{ g C yr}^{-1}$ in biomass and soils in recent years. Over most other areas of the world, agricultural lands have expanded in recent years³⁰, offering no sink for atmospheric CO_2 .

These findings have several implications for understanding recent perturbations of the global carbon cycle that are leading to an increase in atmospheric CO_2 . First, the amount of carbon that can be stored in soils is likely to be only a small fraction of the annual release from fossil fuels ($5.3 \times 10^{15} \text{ g C yr}^{-1}$; ref. 18), even assuming a substantial increase in net primary production as a result of higher atmospheric CO_2 . If the terrestrial biosphere acts as a sink for atmospheric CO_2 , the organic carbon must be found in live-plant biomass and fresh, undecomposed plant litter on the soil surface^{3,31}. Second, recent estimates of the loss of soil organic matter when natural lands are cultivated indicate an annual net release of $0.8 \times 10^{15} \text{ g C}$ (refs 5, 30). Under human influence, the loss of humic materials from cultivated lands is greatly in excess of the rate of formation of humus in undisturbed lands, transforming the soil pool from a small sink to a large source of atmospheric CO_2 . □

Received 1 August; accepted 25 September 1990.

- Schlesinger, W. H. *A. Rev. ecol. Syst.* **8**, 51–81 (1977).
- Post, W. M., Emanuel, W. R., Zinke, P. J. & Stangenberger, A. G. *Nature* **298**, 156–159 (1982).
- Prentice, K. C. & Fung, I. Y. *Nature* **346**, 48–51 (1990).
- Wilson, A. T. *Nature* **273**, 40–41 (1978).
- Schlesinger, W. H. in *The Role of Terrestrial Vegetation in the Global Carbon Cycle* (ed. Woodwell, G. M.) 111–127 (Wiley, New York, 1984).
- Campbell, C. A., Paul, E. A., Rennie, D. A. & McCallum, K. J. *Soil Sci.* **104**, 217–224 (1967).
- Birkeland, P. W. *Soils and geomorphology* (Oxford University Press, 1984).
- Armentano, T. V. & Menges, E. S. *J. Ecol.* **74**, 755–774 (1986).
- Matthews, E. & Fung, I. *Global biogeochem. Cycles* **1**, 61–86 (1987).
- Solomon, A. M., Trabalka, J. T., Reichle, D. E. & Voorhees, L. D. in *Atmospheric Carbon Dioxide and the Global Carbon Cycle* (ed. Trabalka, J. T.) 1–13 (U.S. Department of Energy, Washington, DC, 1985).
- Schlesinger, W. H. *Soil Sci.* **133**, 247–255 (1982).
- Berner, R. A. *Am. J. Sci.* **282**, 451–473 (1982).
- Schlesinger, W. H. & Melack, J. M. *Tellus* **33**, 172–187 (1981).
- Gardner, W. S. & Menzel, D. W. *Geochim. cosmochim. Acta* **38**, 813–822 (1974).
- Flint, R. F. *Glacial and Quaternary Geology* (Wiley, New York, 1971).
- Protz, R., Ross, G. J., Martini, I. P. & Terasmae, J. *Can. J. Soil Sci.* **64**, 31–49 (1984).
- Protz, R., Ross, G. J., Shiptalo, M. J. & Terasmae, J. *Can. J. Soil Sci.* **68**, 287–305 (1988).
- Tans, P. P., Fung, I. Y. & Takahashi, T. *Science* **247**, 1431–1438 (1990).
- Billings, W. D., Luken, J. O., Mortensen, D. A. & Peterson, K. M. *Oecologia* **58**, 286–289 (1982).
- Schleser, G. H. Z. *Naturf.* **37a**, 287–291 (1982).
- Van Cleve, K., Oechel, W. C. & Hom, J. L. *Can. J. Forest Res.* (in the press).
- Emanuel, W. R., Shugart, H. H. & Stevenson, M. P. *Clim. Change* **7**, 29–43 (1985); **7**, 457–460 (1985).
- Prince, A. L., Toth, S. J. & Blair, A. W. *Soil Sci.* **46**, 379–389 (1938).
- Lugo, A. L., Sanchez, M. J. & Brown, S. *Plant Soil* **96**, 185–196 (1986).
- Gebhardt, M. R., Daniel, T. C., Schweizer, E. E. & Allmaras, R. R. *Science* **230**, 625–630 (1985).
- Hart, J. F. *Ann. Ass. Am. Geogr.* **58**, 417–440 (1968).
- Sollins, P., Spycher, G. & Topik, C. *Ecology* **64**, 1273–1282 (1983).
- Syers, J. K., Adams, J. A. & Walker, T. W. J. *Soil Sci.* **21**, 146–153 (1990).
- Delcourt, H. R. & Harris, W. F. *Science* **210**, 321–323 (1980).
- Houghton, R. A. *et al. Tellus* **B39**, 122–139 (1987).
- Schiffman, P. M. & Johnson, W. C. *Can. J. Forest Res.* **19**, 69–78 (1989).
- Evans, L. J. & Cameron, B. H. *Can. J. Soil Sci.* **59**, 203–210 (1979).
- Bockheim, J. G. *Arctic Alp. Res.* **11**, 289–306 (1979).
- Birkeland, P. W. *Arctic Alp. Res.* **10**, 733–747 (1978).
- Chandler, R. *Soil Sci. Soc. Am. Proc.* **7**, 454–459 (1942).
- Tezuka, Y. *Jap. J. Bot.* **17**, 371–402 (1961).
- Bilzi, A. F. & Ciolkosz, E. J. *Soil Sci. Soc. Am. J.* **41**, 122–127 (1977).
- Franzmeir, D. P., Whiteside, E. P. & Mortland, M. M. *Mich. Agric. exp. Station Bull.* **46**, 37–57 (1963).
- Burges, A. & Drover, D. P. *Aust. J. Bot.* **1**, 83–94 (1953).
- Bowman, G. M. *Aust. J. Soil Res.* **27**, 607–628 (1989).
- Bowman, G. M. *Proc. R. Soc. Tasmania* **121**, 75–88 (1987).
- Birkeland, P. W. *Geoderma* **34**, 115–134 (1984).
- Vitousek, P. M., Van Cleve, K., Balakrishnan, N. & Mueller-Dombois, D. *Biotropica* **15**, 268–274 (1983).
- Harris, S. A. in *Paleopedology* (ed. Yaalon, D. H.) 191–209 (Int. Soc. Soil Sci., Jerusalem, 1971).
- Lajtha, K. & Schlesinger, W. H. *Ecology* **69**, 24–39 (1988).
- Crocker, R. L. & Dickson, B. A. *J. Ecol.* **45**, 169–185 (1957).

ACKNOWLEDGEMENTS. I thank D. A. Livingstone, D. W. Johnson, D. S. Schimel, L. D. Schlesinger and B. R. Strain for comments, and G. M. Bowman for unpublished data.

Evidence for chaotic fault interactions in the seismicity of the San Andreas fault and Nankai trough

Jie Huang & D. L. Turcotte

Department of Geological Sciences, Cornell University, Ithaca, New York 14853, USA

INTERACTIONS between fault segments are one of the main sources of complexity in the seismicity of active tectonic regions. Such interactions are likely to influence the spatial and temporal patterns of earthquakes. Here we examine the dynamical behaviour introduced by fault interactions using a simple spring-loaded, slider-block model with velocity-weakening friction. The model consists of two slider blocks coupled to each other and to a constant-velocity driver by elastic springs. For an asymmetric system in which the frictional forces on the two blocks are not equal, the solutions exhibit chaotic behaviour. The system's behaviour over a range of parameter values seems to be generally analogous to that of weakly coupled segments of an active fault. We see similarities between our model simulations and observed patterns of seismicity on the south central San Andreas fault in California and in the Nankai trough along the coast of south-western Japan.

Earthquakes in a zone of crustal deformation are a complex phenomenon. A standard procedure for determining whether a complex system exhibits chaotic behaviour is to study a low-order analogue model. A widely recognized analogue model for seismic faulting is a spring-loaded slider-block system (see refs 1–5, for example). This deterministic mechanical model incorporates the essential elements of the earthquake mechanism, including elastic storage of energy and friction-controlled stick and slip behaviour.

Dynamical instabilities associated with complicated friction laws are well known from studies using single-block models^{6–10}. Under certain conditions, the motion of a system subject to rate- and state-dependent friction laws is characterized by chaotic oscillations⁹. Analyses of the dynamics of a two-block model with spatial symmetry show spatially asymmetric events but no chaotic behaviour¹¹. The introduction of asymmetries, however, results in chaotic behaviour¹². Simulations of homogeneous many-slider systems approximating a continuous fault have shown size distributions that are consistent with the Gutenberg–Richter relation¹³.

Here we model interactions between two weakly coupled fault segments. Our two-block model is illustrated in Fig. 1 and the applicable equations of motion are

$$m_1 \ddot{y}_1 + (k_1 + k_c)y_1 - k_c y_2 = F_1 \quad (1)$$

$$m_2 \ddot{y}_2 + (k_2 + k_c)y_2 - k_c y_1 = F_2 \quad (2)$$

where F_1 and F_2 are the friction forces exerted on masses m_1 and m_2 , which are coupled to each other by a spring with spring constant k_c . The masses are coupled to the slider block by springs of constants k_1 and k_2 , which are extended by distances y_1 and y_2 ($\ddot{y}_i = d^2 y_i / dt^2$). We chose a velocity-weakening friction law of the form

$$F = \frac{F_0}{1 + |\dot{y}|/v_f} \quad (3)$$

where v_f is a reference velocity. The frictional forces decrease monotonically as the sliding velocity increases. Velocity-weakening phenomena have been widely observed in laboratory experiments (refs 14, 15, for example) and this simplified form of velocity-weakening has been used in previous model studies¹³.

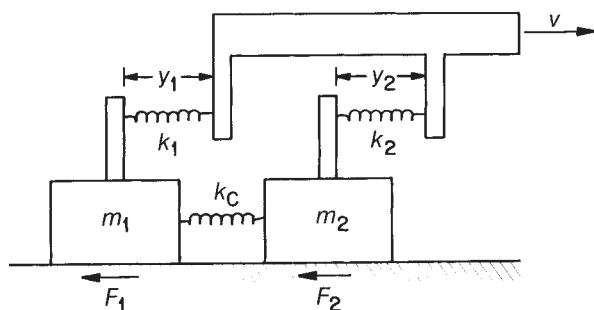


FIG. 1 Illustration of the slider-block fault model. The system is slowly loaded by the driver.

The model is further simplified by introducing the coupling spring constant α , such that $k_c = \alpha k$, and the ratio of the frictional forces of the two blocks β , such that $F_2 = \beta F_1 = \beta F$ (ref. 12). After introducing the dimensionless variables $Y_i = y_i k / F_0$ ($i = 1, 2$) and a dimensionless time $\tau = t \sqrt{k/m}$, the failure criteria on Y_1 and Y_2 are now given by

$$Y_1 + \alpha(Y_1 - Y_2) = 1 \quad (4)$$

$$Y_2 + \alpha(Y_2 - Y_1) = \beta \quad (5)$$

and the equations of motion during slip are

$$\ddot{Y}_1 + Y_1 + \alpha(Y_1 - Y_2) = \frac{1}{1 + \gamma|\dot{Y}_1|} \quad (6)$$

$$\ddot{Y}_2 + Y_2 + \alpha(Y_2 - Y_1) = \frac{\beta}{1 + \gamma|\dot{Y}_2|} \quad (7)$$

where $\gamma = F_0 / v_i \sqrt{mk}$ is a measure of the velocity weakening. The equations are numerically integrated using a fourth-order Runge-Kutta scheme.

With spatial heterogeneity, that is, with the ratio of frictional forces β not equal to one, the dynamical behaviour of the system is generally chaotic (a more detailed analysis of this model will be published elsewhere¹⁶). Figure 2 illustrates a chaotic orbit with 100 events plotted in the phase plane (Y_1, Y_2). Sometimes we observe single-block failures denoted by the vertical and horizontal segments of the orbit, but at other times both blocks slide together (curved segments). The trajectory does not settle into a closed orbit and tends to fill a portion of the phase plane; this is typical of chaotic evolution. Bifurcation analyses reveal that the transitions between stable cyclic behaviour and chaos are characterized by the period-doubling route to chaos, and

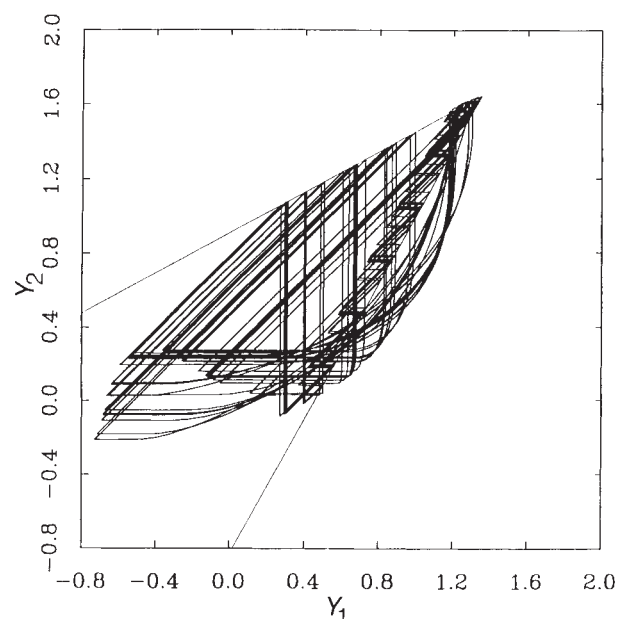


FIG. 2 Space-filling evolution of the system in the Y_1, Y_2 phase plane with system parameters $\alpha = 1.2$, $\beta = 2.0$, $\gamma = 3.0$. The plotted orbit consists of 100 events. The diagonal converging lines are the failure envelope given by equations (4) and (5). Stress accumulation occurs on diagonal lines of unit slope, and failure occurs when the accumulation lines intersect the failure envelope.

the bifurcation values of the period-doubling sequences satisfy the well-known Feigenbaum relation¹⁷.

An important feature in Fig. 2 is that major events involving both blocks are often separated by a sequence of smaller single-block events. The time evolution of the dimensionless cumulative displacement for both slider blocks is given in Fig. 3a. This kind of interaction between the two blocks is clearly displayed at several places where five or six small events of the weaker block occur between major events. When a major event is preceded by a sequence of small events, it will occur along with the last small event in that sequence. The repeat time for the small event before a major event can be much shorter than the average repeat time of the small events and is highly irregular.

This type of behaviour may be representative of the interactions between two weakly coupled fault segments. An interesting example is the interaction between the Parkfield segment and

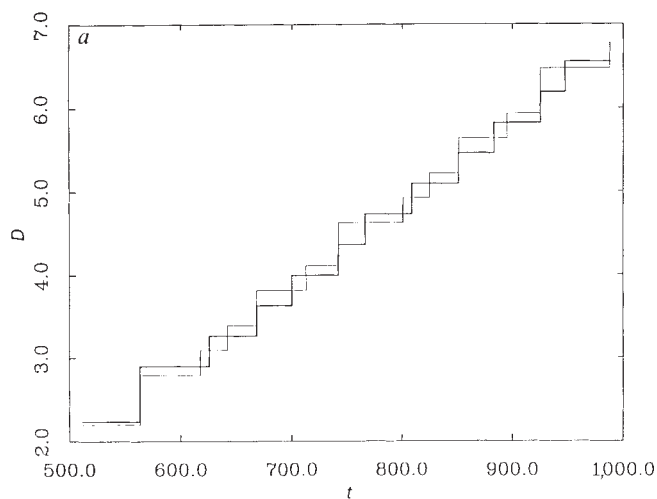
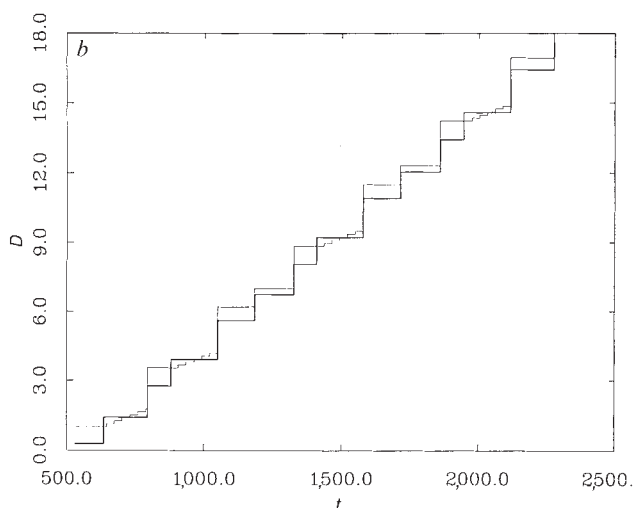


FIG. 3 Time evolution of the dimensionless cumulative displacements for two slider blocks. The lighter trace follows the movement of the weaker block, and the heavy line represents the stronger block. a, System parameters

$\alpha = 1.2$, $\beta = 2.0$, $\gamma = 3.0$, as in Fig. 2. b, System parameters $\alpha = 0.81$, $\beta = 1.05$, $\gamma = 3.0$.

the rest of the south central locked segment of the San Andreas fault. There are strong similarities between the observed seismicity and our model behaviour if we consider the sequence of small events to be analogous to the Parkfield earthquakes in 1881, 1901, 1922, 1934 and 1966, and the big events sandwiching those smaller events to be the 1857-type great earthquakes which ruptured the whole south central section of the San Andreas fault. Using such an analogy, accumulated strain is only partially relieved in each Parkfield event, and the rest will be lost in the coupled major event. The model simulation suggests two scenarios for a great southern California earthquake to occur following a sequence of Parkfield events. In one case, a Parkfield event will transfer sufficient stress to trigger the major event and the major event will have a typical Parkfield timing. It is also possible that, after the last Parkfield event, only a small amount of further strain will cause the major event to occur, carrying the Parkfield block with it. In this case, the major event usually occurs shortly after a Parkfield event.

Quantitative comparisons can be made between the model behaviour and observations. Because our model results are dimensionless, it is convenient to compare the ratios of observables. The two characteristics we compare are the coseismic slip D and the average repeat time T . The coseismic slip for the 1966 Parkfield earthquake is estimated to be $D_s = 0.56$ m (ref. 18), and the mean recurrence time for the Parkfield earthquake sequence is about $T_s = 22$ yr (ref. 19). By measuring the offsets of stream gullies, Sieh²⁰ estimated the coseismic slip for the 1857 great earthquake to be $D_m \approx 8$ m. A mean repeat time of $T_m \approx 132$ yr was derived from the estimates of average long-term slip rate²¹. These estimates give the observed ratio of coseismic slip, $D_m/D_s \approx 14.3$, and the observed ratio of repeat times, $T_m/T_s \approx 6.0$. Our model exhibits a ratio of average displacements between large and small events, $D_m/D_s \approx 14.6$, and the ratio of average repeat times (averaged over 15 cycles of alternating large and sets of small events) is $T_m/T_s \approx 6.0$, in good agreement with the observations.

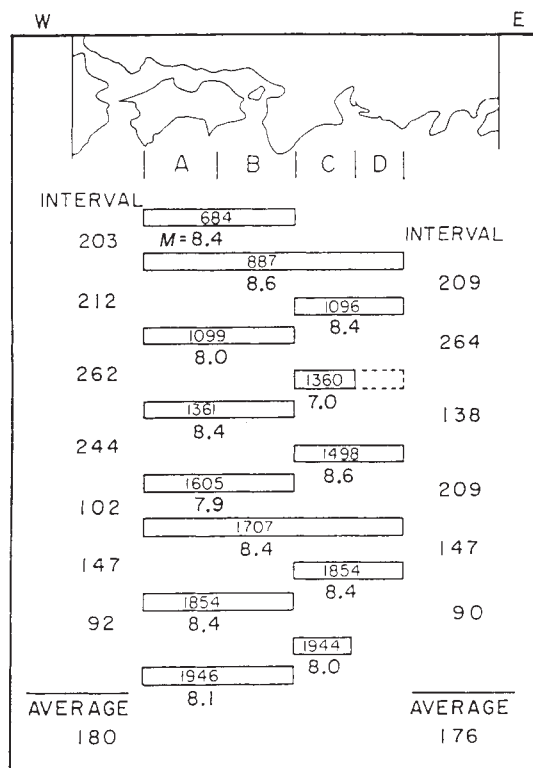


FIG. 4 A sequence of great earthquakes that occurred between 684 and 1946 AD in the Nankai trough along the coast of southwestern Japan (from refs 22, 23).

A variety of fault interactions are also identified at subduction zones. An interaction involving two segments of a subduction zone fault is illustrated by the behaviour of the 200-km-long section of the Nankai trough megathrust along the coast of southwestern Japan. Historical records document a series of great earthquakes that occurred between 684 and 1946 AD, as shown in Fig. 4^{22,23}. The sequence is marked by an irregular but somewhat repetitive pattern in which whole-section failures occur following several alternate failures of single segments.

The simulation of fault interactions with dip-slip motion requires different values of model parameters compared with the previous strike-slip example. Instead of the compression-tension coupling along a strike-slip fault, the form of coupling between the two fault segments is predominantly shear in the case of subduction zones. Hence, the degree of coupling measured by the parameter α should be considerably weaker than that on the San Andreas fault. Moreover, because the earthquakes occurring on both segments are in the same magnitude range, the underlying system should be nearly symmetrical (β close to 1.0). We therefore examined the model evolution with $\alpha = 0.81$, $\beta = 1.05$, $\gamma = 3.0$, a set of parameters which is consistent with the above physical considerations. Figure 3b demonstrates the following sequence of events: a double-block failure, a series of three alternating single-block events, another double-block event, two alternating single-block events, another double-block event, and four more alternating single-block events. We observe this striking similarity between our system behaviour and the observed patterns of seismicity shown in Fig. 4 only for parameter values that are physically reasonable for subduction zones.

As chaotic behaviour of a low-order system often implies chaotic behaviour in similar higher-order systems, our model behaviour suggests that the crustal deformation associated with active faults is likely to be chaotic. The chaotic nature of the seismic faulting process explains the difficulties encountered in the application of exact recurrence models and seismic-gap considerations. Our results show, however, that chaotic behaviour does not exclude the recurrence of certain patterns of seismic activity. The pattern of events from our model simulations indicate that there may be a direct correlation between the occurrence of Parkfield events and a great southern California earthquake. It is generally believed that the short timespan of good-quality seismic records is one of the critical obstacles to the success of earthquake prediction. In this respect, the non-linear dynamics of low-order analogue systems are likely to provide valuable information on patterns of seismicity, which should help in earthquake prediction. □

Received 9 July; accepted 2 October 1990.

- Burridge, R. & Knopoff, L. *Bull. seism. Soc. Am.* **57**, 341–371 (1967).
- Dieterich, J. H. *J. geophys. Res.* **77**, 3771–3781 (1972).
- Rundle, J. B. & Jackson, D. D. *Bull. seism. Soc. Am.* **67**, 1363–1377 (1977).
- Cohen, S. *J. geophys. Res.* **82**, 3781–3796 (1977).
- Cao, T. & Aki, K. *Pure appl. Geophys.* **124**, 487–513 (1986).
- Byerlee, J. *Pure appl. Geophys.* **116**, 616–626 (1978).
- Dieterich, J. H. in *Mechanical Behavior of Crystal Rocks* (eds Carter, N. et al.) Geophys. Monogr. **24**, 103–120 (Am. geophys. Un., Washington, DC, 1981).
- Ruina, A. *J. geophys. Res.* **88**, 10359–10370 (1983).
- Gu, J. et al. *J. Mech. Phys. Solids* **32**, 167–196 (1984).
- Rice, J. R. & Tse, S. T. *J. geophys. Res.* **91**, 512–530 (1986).
- Nussbaum, J. & Ruina, A. *Pure appl. Geophys.* **125**, 629–656 (1987).
- Huang, J. & Turcotte, D. L. *Geophys. Res. Lett.* **17**, 223–226 (1990).
- Carlson, J. M. & Langer, J. S. *Phys. Rev. A* **40**, 6470–6484 (1989).
- Tullis, T. E. & Weeks, J. D. *Pure appl. Geophys.* **124**, 383–414 (1986).
- Shimamoto, T. & Logan, J. M. in *Earthquake Source Mechanics* (ed. Das et al.) Geophys. Monogr. **37**, 49–64 (Am. geophys. Un., Washington, DC, 1986).
- Huang, J. & Turcotte, D. L. *Pure appl. Geophys.* (submitted).
- Feigenbaum, M. J. *J. statist. Phys.* **19**, 25–52 (1978).
- Tsai, Y. B. & Aki, K. *Bull. seism. Soc. Am.* **59**, 275–287 (1969).
- Bakun, W. H. & McEvilly, T. V. *J. geophys. Res.* **89**, 3051–3058 (1984).
- Sieh, K. E. *Bull. seism. Soc. Am.* **68**, 1421–1448 (1978).
- Sieh, K. E. *J. geophys. Res.* **94**, 603–623 (1989).
- Yonekura, N. *Bull. Dept Geogr. Univ. Tokyo* **7**, 19–71 (1975).
- Ando, M. *Tectonophysics* **27**, 119–140 (1975).

ACKNOWLEDGEMENTS. This research was supported by the NASA. Contribution No. 855 from the Department of Geological Sciences, Cornell University.

New *Sivapithecus* humeri from Pakistan and the relationship of *Sivapithecus* and *Pongo*

David Pilbeam*, Michael D. Rose†, John C. Barry* & S. M. Ibrahim Shah‡

* Department of Anthropology, Peabody Museum, Harvard University, Cambridge, Massachusetts 02138, USA

† Department of Anatomy, New Jersey Medical School, 185 South Orange Avenue, Newark, New Jersey 07103-2757, USA

‡ Geological Survey of Pakistan, Quetta, Pakistan

NEW humeri of two species of the Miocene hominoid *Sivapithecus* are described from near Chinji in Pakistan from between ~9 and 11 Myr ago. *Sivapithecus*, a middle and late Miocene hominoid from Turkey and Indo-Pakistan, is overall unlike any living hominoid, although facial-palatal similarities to the extant orangoutan, *Pongo*, have been used to support a hypothesis of close relationship. Living hominoids have postcranial similarities assumed to be shared derived, among them features of the proximal humerus. However, the new *Sivapithecus* proximal humeri differ from those of living hominoids, supporting an alternative hypothesis in which *Sivapithecus* and *Pongo* are not closely related. It is not clear how to choose between these incompatible hypotheses.

The material comes from ~18 to ~1 Myr old Siwalik Group sediments exposed on the Potwar Plateau near Rawalpindi in Pakistan. Collections have been made there since 1973. Some 40,000 specimens including about 160 hominoids are known from over 800 localities¹⁻⁷. Age determinations are based mainly on an extensive palaeomagnetic sampling programme⁸⁻¹¹. The hominoid material is assigned to species of *Sivapithecus*¹², thought to be the sister taxon of *Pongo* on the basis of facial similarities hypothesized as shared derived^{6,13-15}. In other features—occlusal⁶, mandibular¹⁶, postcranial¹⁷⁻¹⁹—*Sivapithecus* and *Pongo* are mostly dissimilar.

GSP 30754 (Figs 1 and 2) is a left humeral shaft lacking epiphyses from locality Y311 in the Nagri Formation, 9.3 Myr old. It is assigned to *S. parvada*²⁰, a newly described species distinct from other *Sivapithecus* species. It is morphologically similar to non-hominoid catarrhines and earlier hominoids, and is clearly not from a large carnivore or other non-primate mammal. A previously described partial humeral distal epiphysis from Y311, GSP 12271 (ref. 6), is similar in preservation and size to, and probably conspecific with, GSP 30754. The proximal third of the specimen curves laterally and anteriorly to its junction with the rest of the shaft. There is a flat, well-defined deltoid plane and associated crests. The olecranon fossa is deep and wide, with articular surface extending onto the lateral wall. GSP 30730, from locality Y76 in the Chinji Formation, 10.8 Myr old, although crushed and somewhat distorted, is an almost complete left humerus, probably of *S. indicus*, lacking only the head (Figs 1 and 2). The proximal shaft is oriented as in GSP 30754. As in GSP 12271 the distal articular surfaces resemble those of modern large hominoids. There is a well-rounded capitulum separated from the lateral trochlear ridge by a well developed groove, the *zona conoidea*, the ridge extending to the lateral wall of the olecranon fossa. The trochlear surface is markedly spool-shaped. Among living primates the distal humeral functional features of GSP 30730 and GSP 12271 are shared only with large hominoids, although the particular combination of features differs from that of any individual hominoid.

Proximal humeral shaft morphology similar to that of the *Sivapithecus* specimens is also found in *Proconsul*²¹, *Kenya-pithecus*¹⁹, and most living quadrupedal non-hominoid primates (Table 1). In suspensory and/or climbing anthropoids, especially hominoids, the shaft is straight and bears a convex deltoid plane^{19,21,22}. The elbow joint complex in living hominoids provides stabilization throughout extensive flexion-extension and pronation-supination excursions²³. GSP 30730 is probably the oldest known hominoid with the distal humeral morphology found in modern large hominoids, yet it is unusual in combining this with a curved proximal shaft morphology. (The distal humeri of *Rudapithecus*²⁴ and *Oreopithecus*^{25,26} also

FIG. 1 Anterior views of humeri of *a*, *Papio hamadryas* (right). *b*, *Pan troglodytes* (right). *c*, *Pongo pygmaeus* (right). *d*, GSP 30754 (left). *e*, GSP 12271 (cast, right). *f*, GSP 30730 (left). Scale bar (cm).

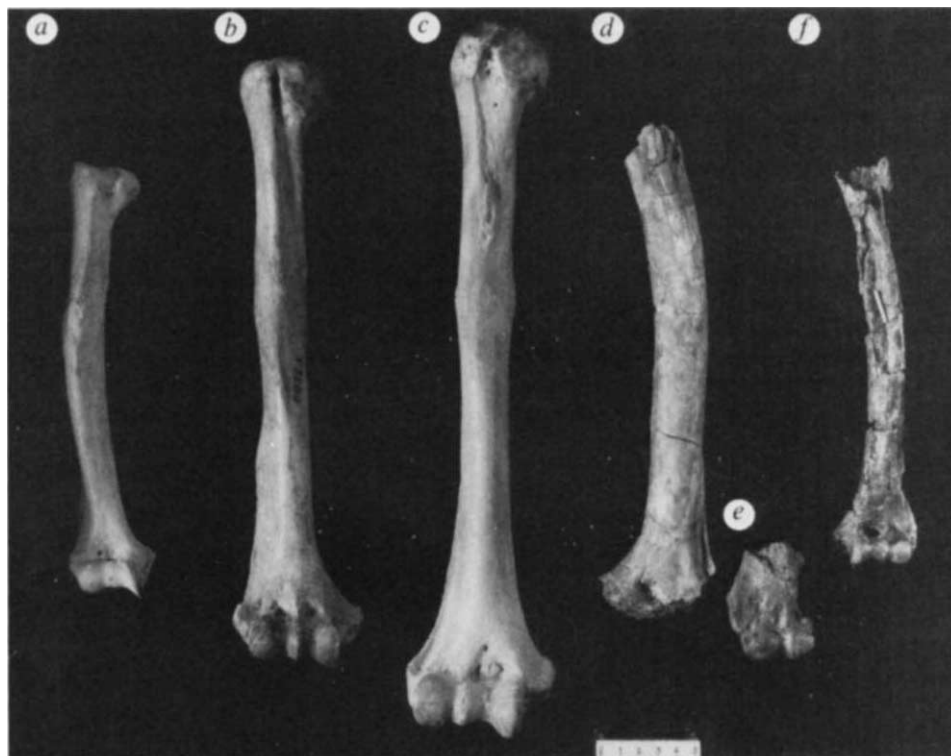


TABLE 1 Distribution of some humeral features

	<i>Proconsul</i>	<i>Kenyapithecus</i>	<i>Sivapithecus</i>	<i>Rudapithecus</i>	<i>Cercopithecids</i>	African apes	<i>Pongo</i>
Medially inclined proximal shaft	?	?	+	?	+	—	—
Retroflexed proximal shaft	+	+	+	?	+	—	—
Flat deltoid plane	+	+	+	?	+	—	—
Mediolaterally broad trochlea	+	+	+	+	—	+	+
Markedly spoon-shaped trochlea	—	—	+	+	—	+	+
Prominent lateral trochlear keel and deep, narrow <i>zona conoidea</i>	—	—	+	+	—	+	—

? = Not known; + = Feature present; — = Feature absent.

show a derived hominoid condition, but their shaft morphology is not known, or is unclear because of crushing.) The functional pattern of the hominoid distal humerus has generally been interpreted in terms of suspensory behaviour, with variations in the pattern in living hominoids being linked to particular mixtures of quadrupedalism, climbing and suspension in individual species^{27–29}. Complete and partial *Sivapithecus* postcranial specimens are known from all parts of the limbs except the leg^{6,17–19}. These, together with the new specimens, indicate that *Sivapithecus* was, like *Proconsul* and *Kenyapithecus*, and perhaps other more poorly known Miocene hominoids, basically quadrupedal and that climbing and suspension were probably less emphasized in the locomotor repertoire. This evidence suggests that features, including those of the distal humerus, now associated with specialized suspensory and climbing functions in hominoids had their origin in an already partly derived morphology that was first associated with more quadrupedal and less suspensory activities.

Pongo and *Sivapithecus* are similar in certain palatal and facial characteristics by which they differ from all other living and fossil hominoids, and which have been interpreted as shared derived characteristics^{13–15}. There are no such features unequivocally present in the known parts of the *Sivapithecus* postcranium.

As mentioned above for the distal humerus, *Sivapithecus* postcranials do share several features with extant large hominoids, particularly African apes. Thus the calcaneo-cuboid joint includes features related to extensive, stabilized pronation-supination movements¹⁸, a stable semi-flexed and medially rotated position of the external pollex is indicated by proximal phalangeal features, and the hallux is robust, with a broad flat terminal phalanx⁶. These features have been interpreted as being primitive for large hominoids¹⁹. The above-mentioned humeral shaft features, together with features of the talus, medial calcaneus, distal cuboid, proximal capitate and the proximal manual phalangeal base are primitive with respect to the apparently shared derived state of these features in living large hominoids. Functionally, these primitive features are associated with generalized quadrupedalism and are shared by *Sivapithecus*, more primitive fossil hominoids, and many living non-hominoid anthropoids. This is not to be expected if *Sivapithecus* is the sister taxon of *Pongo*.

Two mutually exclusive hypotheses follow. First, that *Sivapithecus* and *Pongo* are sister taxa, in which case a number of postcranial features shared by living large hominoids must represent convergences. Second, that *Sivapithecus* and *Pongo* are not sister taxa, in which case their palatal and facial

FIG. 2 Lateral views of humeri of *a*, *Papio hamadryas* (right). *b*, *Pan troglodytes* (right). *c*, *Pongo pygmaeus* (right). *d*, GSP 30754 (left). *e*, GSP 12271 (cast, right). *f*, GSP 30730 (left). Scale bar (cm).



similarities are not shared derived features but either convergent derived or shared primitive features. (A variant of the first hypothesis, that *Sivapithecus* and *Pongo* are sister taxa, and the apparently primitive features of the *Sivapithecus* postcranium are reversals from a more derived condition, is unlikely in light of the close similarity of these features in *Sivapithecus* and the taxa mentioned above.) Distinguishing between the two hypotheses will require detailed studies of apparently shared derived postcranial features of living hominoids, and a re-examination of the apparently shared derived palatal and facial features of *Sivapithecus* and *Pongo*. There is some evidence that

marked torsion of the humeral head (not determinable for any currently available fossil hominoid) may have been acquired independently in at least some living hominoid lineages^{19,30,31}. The hominoids are a group of low taxonomic diversity and high morphological variability. But even with a denser hominoid fossil record, problems will remain with the objective definition of characters, assessment of homology versus convergence as alternative explanations for similarity of character states³², and determination of whether character states are primitive or derived. We are not confident that biologically plausible procedures exist for unambiguously settling these issues. □

Received 14 May; accepted 27 September 1990.

1. Pilbeam, D. et al. *Nature* **270**, 684–689 (1977).
2. Pilbeam, D. et al. *Nature* **270**, 689–695 (1977).
3. Barry, J. C., Johnson, N. M., Raza, S. M. & Jacobs, L. L. *Geology* **13**, 637–640 (1985).
4. Behrensmeyer, A. K. *Soc. Ec. Paleont. Mineral. Sp. Pub. no. 39* (1987).
5. Quade, J., Cerling, T. E. & Bowman, J. R. *Nature* **342**, 163–166 (1989).
6. Pilbeam, D. R., Rose, M. D., Badgley, C. & Lipschutz, B. *Postilla* (Yale Peabody Museum, 1980).
7. Raza, S. M. et al. *Nature* **306**, 52–54 (1983).
8. Johnson, N. M., Stix, J., Tauxe, L., Cervený, P. F. & Tahirkehi, J. *Geology* **93**, 27–40 (1985).
9. Kappelman, J. thesis, Harvard Univ. (1986).
10. Johnson, H. M., Sheikh, K. A., Dawson-Saunders, E. & McRae, L. E. in *New Perspectives in Basin Analysis* (eds Kleinspehn, K. L. & Paola, C.) 189–200 (Springer, New York, 1988).
11. Berggren, W. A., Kent, D. V., Flynn, L. J. & Van Couvering, J. A. *Geol. Soc. Am. Bull.* **96**, 1407–1418 (1985).
12. Pilgrim, G. E. *Rec. geol. Surv. India* **40**, 63–71 (1910).
13. Andrews, P. & Cronin, J. E. *Nature* **297**, 541–546 (1982).
14. Ward, S. C. & Kimbel, W. H. *Amer. J. phys. Anthropol.* **61**, 157–171 (1983).
15. Ward, S. C. & Pilbeam, D. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 211–238 (Plenum, New York, 1983).
16. Brown, B. thesis, Kent State Univ. (1989).
17. Rose, M. D. *J. hum. Evol.* **13**, 503–516 (1984).

18. Rose, M. D. *J. hum. Evol.* **15**, 333–367 (1986).
19. Rose, M. D. *J. hum. Evol.* **18**, 131–162 (1989).
20. Kelley, J. J. *J. hum. Evol.* **17**, 305–324 (1988).
21. Napier, J. R. & Davis, P. R. *Fossil Mammals Afr.* **16**, 1–69 (1959).
22. Erikson, G. E. in *The Primates* (eds Napier, J. & Barnicot, N. A.) 135–163 (Zoological Society of London, 1963).
23. Rose, M. D. *J. hum. Evol.* **17**, 193–224 (1988).
24. Morbeck, M. E. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 369–404 (Plenum, New York, 1983).
25. Harrison, T. *J. hum. Evol.* **15**, 541–583 (1986).
26. Sarmiento, E. E. *Novitates* (American Museum of Natural History, 1987).
27. Keith, A. J. *Anat. Physiol.* **37**, 18–40 (1903).
28. Feldesman, M. R. *Am. J. phys. Anthropol.* **59**, 73–95 (1982).
29. Senut, B. in *Primate Evolution Vol. 1* (eds Else, J. G. & Lee, P. C.) 151–161 (Cambridge University Press, 1986).
30. Sarmiento, E. E. thesis, New York Univ. (1985).
31. Larson, S. G. *Am. J. phys. Anthropol.* **76**, 449–462 (1988).
32. Simpson, G. G. *Principles of Animal Taxonomy* (Columbia University Press, New York, 1961).

ACKNOWLEDGEMENTS. This work was supported by the NSF and SFOP (D.P. and J.C.B.); and the NIH and NSF (M.D.R.).

Activation of facilitation calcium channels in chromaffin cells by D₁ dopamine receptors through a cAMP/protein kinase A-dependent mechanism

Cristina R. Artalejo*, Marjorie A. Ariano†, Robert L. Perlman*‡ & Aaron P. Fox*§

The University of Chicago * Department of Pharmacological and Physiological Sciences, † Department of Pediatrics and the Joseph P. Kennedy Jr Mental Retardation Research Center, 947 E. 58th St, Chicago, Illinois 60637, USA

‡ Department of Anatomy and Neurobiology, University of Vermont, Burlington, Vermont 05405, USA

FACILITATION calcium channels^{1–4} in unstimulated bovine chromaffin cells are normally quiescent³ but are activated by large pre-depolarizations or by repetitive depolarization in the physiological range. The activation of these 27-pS dihydropyridine-sensitive channels by repetitive stimulation^{4,5}, such as by increased splanchnic nerve activity, can lead to an almost twofold increase in Ca²⁺ current in these cells³. This increase in Ca²⁺ current is of probable physiological importance in stimulating rapid catecholamine secretion in response to danger or stress. We have identified D₁ dopaminergic receptors on bovine chromaffin cells by fluorescence microscopy⁶. Here we show that stimulation of the D₁ receptors activates the facilitation Ca²⁺ currents in the absence of pre-depolarizations or repetitive activity, and that activation by D₁ agonists is mediated by cyclic AMP and protein kinase A. The recruitment of facilitation Ca²⁺ channels by dopamine may form

the basis of a positive feedback loop mechanism for catecholamine secretion.

Figure 1a, a plot of peak current as a function of test pulse potential, shows that the dopamine agonist apomorphine (0.1 µM) greatly increases the Ca²⁺ current. Nisoldipine (1 µM), a dihydropyridine antagonist, completely inhibited the Ca²⁺ current recruited by apomorphine (Fig. 1a). It has been shown previously that dihydropyridine antagonists such as nisoldipine specifically suppress the facilitation Ca²⁺ current but do not affect other currents in bovine chromaffin cells³; thus, apomorphine seems to activate only the facilitation Ca²⁺ current. Figure 1b, a plot of peak current versus time, illustrates that dopamine (10 µM) could also recruit the facilitation Ca²⁺ current. The effects of dopaminergic agonists were apparently mediated by D₁ dopamine receptors as the specific D₁ dopamine agonist SKF-38393 (1 µM) activated the facilitation Ca²⁺ current (Fig. 1c). In 11 experiments the D₁ agonist increased chromaffin cell Ca²⁺ current to 182 ± 6% of control. The D₁ response could be prevented by the specific D₁ antagonist SCH 23390 (1 µM; Fig. 1c) but not by the serotonergic antagonist ketanserin (10 µM, N = 4; not shown). The D₂ agonist quinpirole did not activate facilitation, nor did it have any other effects on Ca²⁺ currents^{7–9}, even at concentrations up to 100 µM (not shown). Large pre-pulses, which activated facilitation Ca²⁺ channels in untreated cells, did not increase Ca²⁺ currents in cells that had been treated with SKF-38393 (Fig. 1d), indicating that pre-pulses and dopamine activated the same group of Ca²⁺ channels. Taken together, these results suggest that the activation of D₁ dopamine receptors specifically recruited facilitation Ca²⁺ channels in chromaffin cells.

Chromaffin cells have been reported to have D₂ but not D₁ dopamine receptors^{7–9}. Figure 2 confirms the presence of D₁ dopamine receptors in bovine chromaffin cells. Fluorescence microscopy using the rhodamine conjugate of the 4'-amino derivative of the D₁ antagonist SCH 23390 (ref. 6) demonstrated binding to almost all the cells (Fig. 2a). Binding of the fluorescent ligand was not affected by ketanserin (10 µM, Fig. 2b)

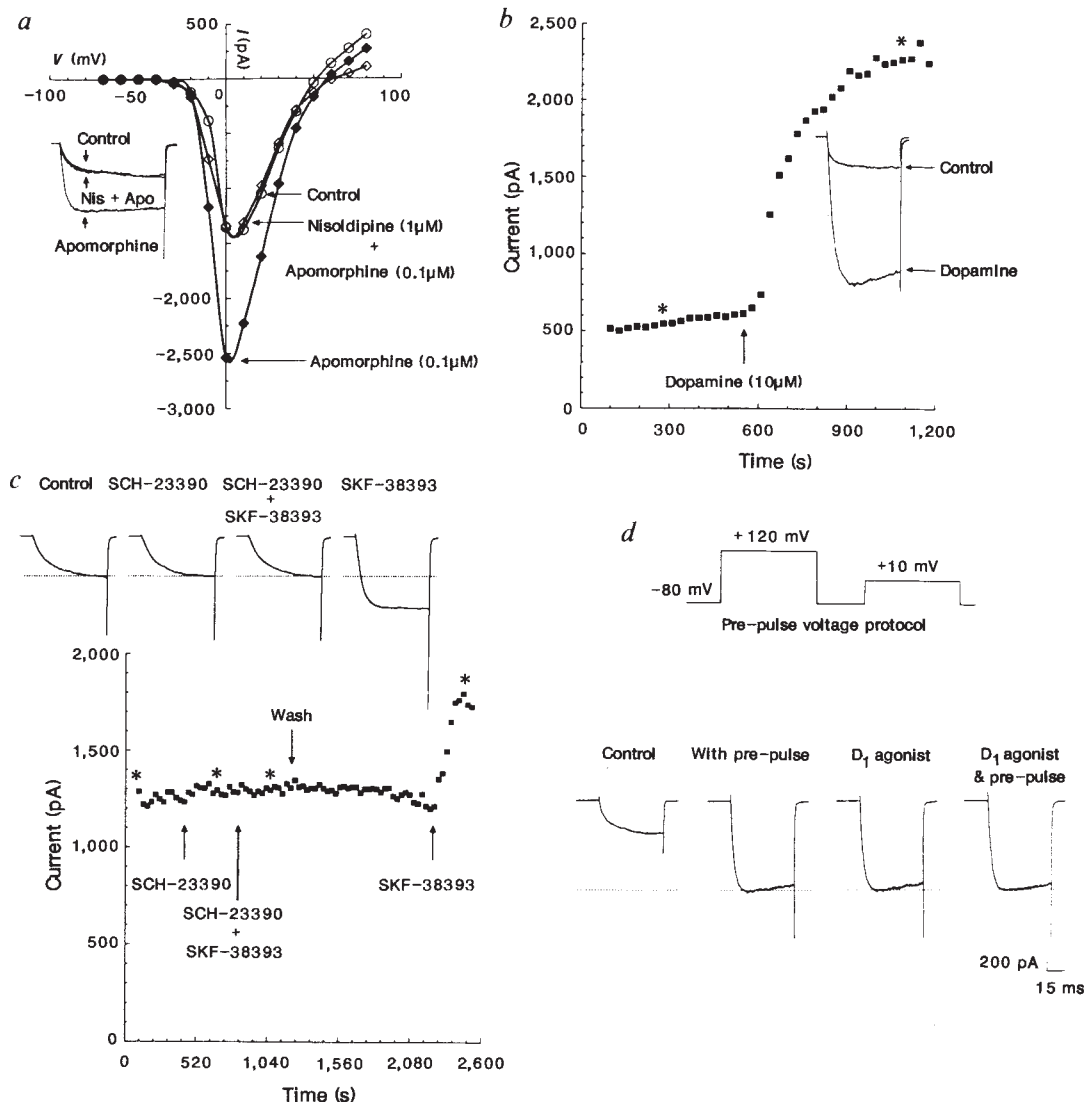
§ To whom correspondence should be addressed.

indicating that the ligand was not binding to serotonin receptors. Nonspecific binding was low (Fig. 2c). Thus, the rhodamine-SCH 23390 conjugate seems to bind specifically to D₁ receptors on chromaffin cells. Using the rhodamine conjugate of the D₂ antagonist *N*-(*p*-aminophenethyl)spiperone, we have confirmed that chromaffin cells also contain D₂ receptors. Companion experiments with both fluorescent ligands demonstrated that the cells had more D₂ than D₁ receptors. The failure of other investigators to detect D₁ receptors on chromaffin cells by ligand-binding techniques may reflect the relatively low density of these receptors on the cells.

In sympathetic neurons¹⁰ and other tissues, D₁ receptors are positively coupled to adenylyl cyclase¹¹. Membrane-permeant cAMP derivatives, 8-(4-chlorophenylthio)-cAMP (8-CPT-cAMP) and 8-Br-cAMP, mimicked the effect of D₁ agonists on the facilitation Ca²⁺ currents. 8-CPT-cAMP (50 μ M) increased the Ca²⁺ current amplitude to 180% $\pm$ 6% of control (N = 13; Fig. 3a). Increasing 8-CPT-cAMP concentrations above 50 μ M did not increase the response amplitude. The responses produced by D₁ receptor activation (SKF-38393) or by 8-CPT-cAMP were not additive (Fig. 3b), suggesting that both increased Ca²⁺ currents by the same mechanism. Inhibition of protein

FIG. 1 Stimulation of D₁ but not D₂ dopamine receptors activates facilitation Ca²⁺ currents. *a*, Plot of peak current versus potential (*I*-*V*) obtained in drug-free conditions (○), in the presence of 0.1 μ M apomorphine (◆), and with both 1 μ M nisoldipine and 0.1 μ M apomorphine (◇). Inset, superimposed peak current traces obtained at +10 mV, with conditions as listed. Test potentials (TP) lasted 50 ms, from a holding potential (HP) of -80 mV. Linear leak and capacitance have been subtracted with 600 μ s blanked at both the on and off of the depolarization. *b*, Plot of peak current versus time (*I*-*T*) in absence and presence of dopamine. The cell was depolarized to +10 mV from HP = -80 mV once every 30 s, conditions repeated in all current-time plots. At the arrow, 10 μ M dopamine was added to the bath. Inset, current records obtained at the point labelled *. *c*, *I*-*T* plot showing that D₁ agonists could activate facilitation Ca²⁺ currents whereas D₁ antagonists could prevent the D₁ agonist response. At the first arrow 1 μ M of the D₁ antagonist SCH-23390 was added to the bath. At the second arrow SKF-38393 was added in the continued presence of SCH-23390. At the third arrow the D₁ agonist and antagonist were washed out of the bath. At the fourth arrow the D₁ agonist was reapplied, showing that in the absence of D₁ antagonist the D₁ agonist could recruit facilitation. Inset, current records obtained at the point labelled *. *d*, D₁ agonists recruited the same population of Ca channels as did large pre-pulses. The current labelled Control was elicited by a depolarization to +10 mV from HP = -80 mV. Large pre-pulses could induce facilitation³. The current labelled 'With pre-pulse' was obtained using the voltage protocol shown on top. Only current from the test pulse, to +10 mV, was plotted. Application of the D₁ agonist also greatly increased the Ca current above control levels, with no pre-pulse applied. When large pre-pulses were applied in the presence of the D₁ agonist, no further increase in current amplitude was observed.

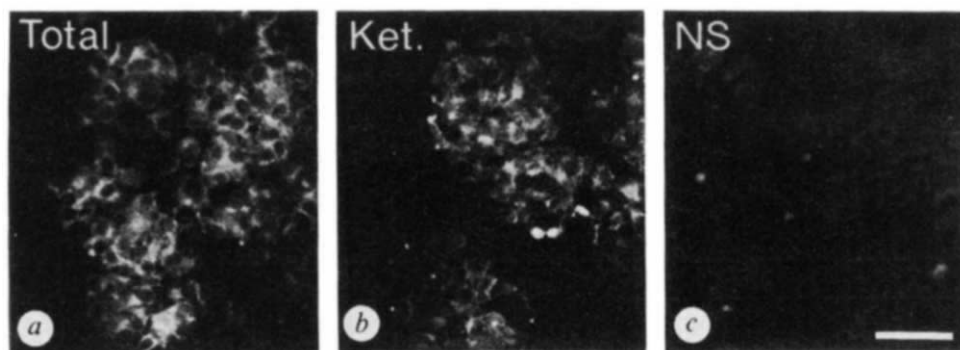
METHODS. Chromaffin cells were isolated from bovine adrenal glands by a modification of the procedure of Greenberg and Zinder¹⁶. Cells were cultured



at 1.5×10^5 cells cm^{-2} in serum-containing medium as previously described. Cells were used 1–10 days after preparation. Whole-cell conditions: solutions were used to suppress Na⁺ and K⁺ channel currents. The internal (pipette) solution consisted of (mM): 100 CsCl, 10 Cs-EGTA, 40 HEPES buffer, 5 MgCl₂, 2 ATP, 0.3 GTP pH 7.3 (with CsOH). The external solution was exchanged after gigaseal formation from a Ringers-type solution to one containing (mM): 10 BaCl₂, 135 triethylammonium chloride, 10 HEPES and 1 μ M tetrodotoxin pH 7.3 (with triethylammonium hydroxide). Voltage pulses and current traces were simultaneously generated and acquired with an IBM PC/AT clone computer. Whole-cell currents were filtered with an 8 pole Bessel filter with corner frequency at 3 KHz and a sampling rate of 5 KHz. Whole-cell stimulation rates were between once every 20 to once every 60 s. All experiments were carried out at room temperature (21–23 °C).

FIG. 2 Bovine chromaffin cells displayed D_1 dopamine receptors. *a*, Fluorescence photomicrograph of total binding of the rhodamine conjugate of the 4'-amino derivative of the D_1 antagonist SCH 23390 (100 nM; see methods below). *b*, D_1 receptor binding was not significantly affected by the serotonin antagonist ketanserin (10 μ M). *c*, Nonspecific binding was determined by inclusion of 10 μ M of the unlabelled antagonist. Calibration bar, 50 μ m.

METHODS. Chromaffin cells were plated on polylysine-coated cover glasses, cultured for 3 days, washed with Earle's Balanced Salt Solution and allowed to air dry. The cells were then incubated overnight at 4 °C in PBS containing the rhodamine-SCH 23390 conjugate (Molecular Probes 100 nM) to determine total D_1 receptor binding. The contribution of serotonin receptors to rhodamine-SCH 23390 binding



was evaluated by measuring binding in the presence of 10 μ M ketanserin. Nonspecific binding was determined in the presence of 10 μ M non-conjugated SCH 23390. Fluorescence of the bound rhodamine-SCH 23390 was observed as described⁶.

kinase A by the specific protein kinase inhibitor peptide (PKI 14-24 amide) suppressed the response to SKF-38393 (Fig. 3c). The PKI peptide (10 μ M) was in the patch-pipette throughout the experiment; to ensure that adequate inhibitor had diffused into the cell, approximately 20 min were allowed to pass between breaking into the cell and starting the experiment. Addition of SKF-38393 produced no effect (Fig. 3c). BAY K 8644 has previously been shown to activate facilitation Ca^{2+} channels directly³. The addition of BAY K 8644 recruited the facilitation Ca^{2+} current indicating that the PKI peptide did not block the channels directly. In eight experiments with PKI in the pipette internal solution, no responses to SKF-38393 were observed; in 11 experiments without the inhibitor, all the cells responded to the SKF-38393. Calmodulin-dependent protein kinase substrate (CDPKS; Peninsula Labs) is a peptide of 10 amino acids similar to PKI but which does not block protein kinase A, and did not inhibit the recruitment of facilitation currents by SKF-38393 ($N=4$). The effects of D_1 receptor activation therefore seem to

be mediated by cAMP and protein kinase A (Fig. 3). The site of protein kinase A phosphorylation is unknown.

Both SKF-38393 (100 μ M) and 8-CPT-cAMP (500 μ M) could increase the unitary activity of the 27-pS facilitation Ca^{2+} channel^{4,5}. Figure 4a (left panel) shows traces obtained under drug-free conditions (control) from a cell-attached patch. Immediately below the current traces are plotted the ensemble currents. The right panel of Fig. 4a illustrates the greatly increased mean open times and probability of channel openings resulting from the addition of the D_1 agonist. Similar dramatic increases in activity of the 27-pS Ca^{2+} channel were observed in the presence of 8-CPT-cAMP (Fig. 4b).

Recruitment of the facilitation Ca^{2+} current by dopamine and cAMP may be an important regulatory mechanism in chromaffin cells. Dopamine, an intermediate in the biosynthesis of noradrenaline and adrenaline, may be secreted together with the other catecholamines⁷. Alternatively, dopamine receptors on chromaffin cells may be stimulated by plasma dopamine. If

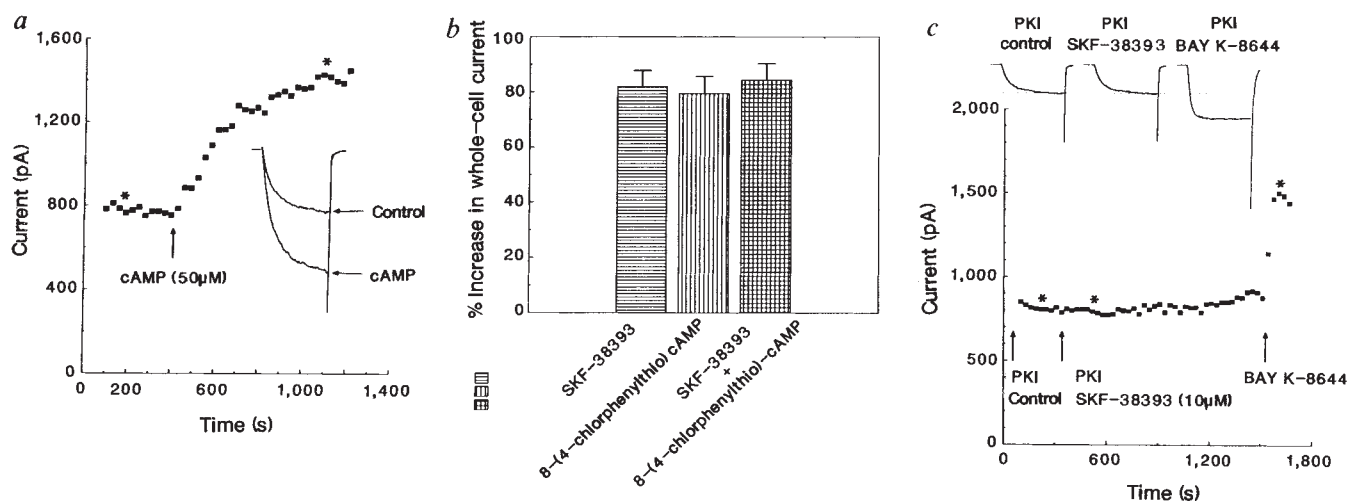
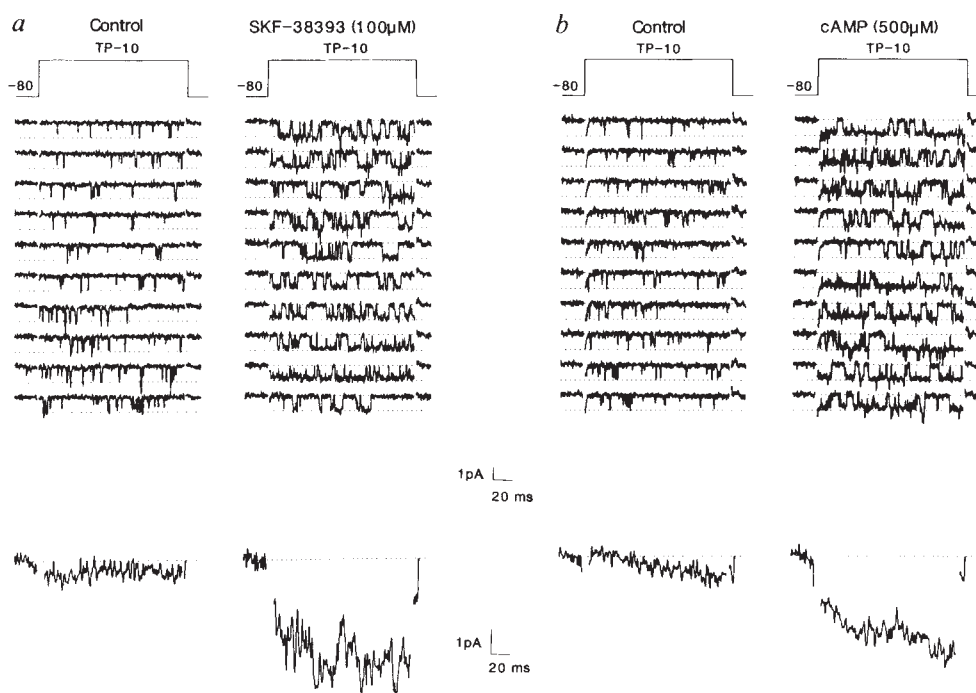


FIG. 3 cAMP analogues mimicked the effect of D_1 agonists, whereas protein kinase A blockers suppressed the D_1 response. *a*, I - T plot illustrating the dramatic increase in whole-cell current produced by the membrane permeant cAMP analogue, 8-CPT-cAMP. At the arrow 50 μ M 8-CPT-cAMP was added to the bath (labelled cAMP (50 μ M)). Inset, current records obtained at the point labelled *. Currents in this experiment were produced by 30 ms depolarizations from HP = -80 mV, to TP = +10 mV. *b*, Cumulated increases

in whole-cell currents produced by SKF-38393 ($N=11$), 8-CPT-cAMP ($N=13$) and the combination of SKF-38393 and 8-CPT-cAMP ($N=8$). *c*, The specific protein kinase A blocker, PKI 14-24 Amide (10 μ M; Peninsula Labs) an 11-amino-acid peptide, could suppress the D_1 response. The PKI was present in the pipette throughout the experiment. At the second arrow, SKF-38393 (10 μ M) was added to the bath. At the third arrow, BAY K 8644 (1 μ M) was introduced into the bath.

FIG. 4 The 27-pS Ca^{2+} channel responsible for facilitation exhibited a greatly increased open probability in the presence of either a D_1 agonist or cAMP. **a**, Representative 27-pS channel activity consecutively recorded in a cell-attached patch in the absence of SKF-38393 (control; left panel), and in the presence of 100 μM SKF-38393 (right panel). All traces have been corrected for linear leak and capacitance. Average currents shown below were obtained by averaging over 200 unitary current records like those shown. These patches were chosen because they exhibited predominantly 27-pS Ca^{2+} channel openings. To see openings of the normally quiescent 27-pS channel under control conditions, cells were depolarized every 4 s, a rate sufficient to recruit some facilitation. **b**, 27-pS channel recorded in the absence of 8-CPT-cAMP (Control, left panel) and in the presence of 500 μM 8-CPT-cAMP (cAMP, right panel). The concentrations of drugs used for these single-channel experiments were higher than those required for the whole-cell recordings. Lowering the drug concentrations introduced a long lag period before a response was observed. As our cell-attached patches had a finite lifetime, we sped up the channel responses by increasing the agonist concentrations. Single-channel conditions: Patch pipette contained (mM): 90 BaCl_2 , 15 CsCl , 15 TEA-Cl , 10 HEPES , pH 7.3 (with CsOH). The cell



resting potential was zeroed using an external solution containing (mM): 140 K-aspartate, 10 K-EGTA, 10 HEPES , 1 mM MgCl_2 pH 7.4 (with KOH). Current signals were sampled at 5 KHz and filtered at 1 KHz. Responses to either cAMP or SKF-38393 were observed in more than 15 patches.

dopamine release from chromaffin cells activates D_1 receptors, the D_1 receptor $\rightarrow$ adenylyl cyclase $\rightarrow$ protein kinase A $\rightarrow$ facilitation Ca^{2+} channel cascade may form a positive feedback loop that augments catecholamine secretion. Other agents that increase intracellular cAMP levels may also recruit facilitation Ca^{2+} channels and enhance Ca^{2+} -dependent processes in the cells. Compounds that raise cAMP levels in chromaffin cells, including vasoactive intestinal peptide¹², adenosine¹³ and forskolin^{14,15}, have all been reported to augment catecholamine secretion. Recruitment of facilitation Ca^{2+} channels may underlie the actions of these agents.

Note added in proof: Recent studies have shown that D_1 receptors can couple to phospholipase C (ref. 17), thereby activating protein kinase C. In four experiments we found activation of protein kinase C by phorbol dibutyrate (100 nM) did not affect facilitation Ca^{2+} currents. □

Received 26 June; accepted 11 September 1990.

1. Fenwick, E. M., Marty, A. & Neher, E. *J. Physiol., Lond.* **331**, 599–635 (1982).
2. Hoshi, T., Rothlein, J. & Smith, S. *J. Proc. natn. Acad. Sci. U.S.A.* **81**, 5871–5875 (1984).
3. Artalejo, C. R., Dahmer, M. K., Perlman, R. L. & Fox, A. P. *J. Physiol., Lond.* (in the press).
4. Hoshi, T. & Smith, S. *J. Neurosci.* **7**, 571–580 (1987).
5. Artalejo, C. R., Dahmer, M. K., Perlman, R. L. & Fox, A. P. *Soc. Neurosci. (Abstr.)* **15**, 651 (1989).
6. Ariano, M. A., Monsma, F. J., Jr, Barton, A. C., Kang, H. C., Haugland, R. P. & Sibley, D. R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8570–8574 (1989).
7. Artalejo, A. R., Garcia, A. G., Montiel, C. & Sanchez, G. P. *J. Physiol., Lond.* **362**, 359–368 (1985).
8. Bigornia, L., Suozzo, M., Ryan, K. A., Napp, D. & Schneider, A. S. *J. Neurochem.* **51**, 999–1006 (1988).
9. Quirk, M., Bergeron, L., Mount, H. & Philie, J. *Biochem. Pharm.* **36**, 3707–3713 (1987).
10. Libet, B. in *Neurotransmitters and Cortical Function. From Molecules to Mind* (eds Avoli, M., Reader, T. A., Dykes, R. W. & Gloor, P. 453–469 (Plenum, New York, 1987).
11. Kebabian, J. W. & Calne, D. B. *Nature* **277**, 93–96 (1979).
12. Wilson, S. P. *Neuropeptides* **11**, 17–21 (1988).
13. Chern, Y.-J., Kim, K. T., Slakey, L. L. & Westhead, E. W. *J. Neurochem.* **50**, 1484–1493 (1988).
14. Morita, K., Dohi, T., Kitayama, S., Koyama, Y. & Tsujimoto, A. *J. Neurochem.* **48**, 243–247 (1987).
15. Marriott, D., Adams, M. & Boarder, M. R. *J. Neurochem.* **50**, 616–623 (1988).
16. Greenberg, A. & Zinder, O. *Cell Tiss. Res.* **226**, 655–665 (1982).
17. Mahan, L. C., Burch, R. M., Monsma, F. J., Jr & Sibley, D. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2196–2200 (1990).

ACKNOWLEDGEMENTS. This work was supported by the NIH, the Klingenstein Fund, the McKnight Foundation, the Whitaker Foundation, the Sprague Foundation, the NSF and the Foundation Juan March. We thank M. Hart for the preparation of cells, C. Palfrey for comments on the manuscript and the PKI peptide, and A. Scribaine and A. Garcia for gifts of nisoldipine and BAY K 8644.

Modulation of glycine receptor chloride channels by cAMP-dependent protein kinase in spinal trigeminal neurons

Yumin Song & Li-Yen Mae Huang*

Marine Biomedical Institute and Department of Physiology and Biophysics, University of Texas Medical Branch, 200 University Blvd, Galveston, Texas 77550, USA

GLYCINE is an important inhibitory transmitter in the brainstem and spinal cord^{1,2}. In the trigeminal subnucleus caudalis (medullary dorsal horn) and in the spinal dorsal horn (the relaying centres for processing pain and sensory information^{3–5}), glycine inhibits the glutamate-evoked depolarization and depresses firing of neurons⁶. The binding of glycine to its receptor produces a large increase in Cl^- conductance, which causes membrane hyperpolarization⁷. The selectivity and gating properties of glycine receptor channels have been well characterized⁸; the glycine receptor molecules have also been purified^{9–11}. The amino-acid sequence, deduced from complementary DNA clones encoding one of the peptides (the 48K subunit)¹², shows significant homology with γ -aminobutyric acid A (GABA_A) and nicotinic acetylcholine receptor subunits, suggesting that glycine receptors may belong to a superfamily of chemically gated channel proteins^{12,13}. However, very little is known about the modulation of glycine receptor channels. We have investigated the regulation of strychnine-sensitive glycine receptor channels by cyclic AMP-dependent protein kinase in neurons isolated from spinal trigeminal nucleus of rat and report here that the protein kinase A dramatically increased the glycine-induced Cl^- currents by increasing the proba-

* To whom correspondence should be addressed.

bility of the channel openings. G_s protein, which is sensitive to cholera toxin, was involved in the modulation.

Glycine elicited large Cl^- currents in 81 of the 85 spinal trigeminal neurons we tested. Application of glycine at the holding potential of -80 mV, under our experimental conditions ($[\text{Cl}^-]_o = 147$ mM; $[\text{Cl}^-]_i = 72$ mM), produced large inward cur-

rents that desensitized slowly (Fig. 1a). The currents were eliminated by $10 \mu\text{M}$ strychnine or by removing Cl^- from the external medium (not illustrated). When $500 \mu\text{M}$ cAMP was internally perfused into the cells, the currents increased to about twice the amplitude (Fig. 1a). The kinetics of the responses did not change. The time course of the change in peak glycine responses is given in Fig. 1a. The enhancing effect of cAMP took 1–2 min to appear, 5 min to reach a stable level and remained stable or sometimes continued to increase for a period of 20–25 min.

This effect of cAMP was quite specific to the glycine response. Intracellular cAMP had no effect on kainate responses in the same cell (Fig. 1b). The responses to GABA in these neurons usually decreased with time, presumably as a result of rundown¹⁴. After switching to a cAMP-containing pipette solution, the GABA-induced Cl^- current continually decreased (Fig. 1c).

Further evidence that cAMP-dependent phosphorylation caused the increase of glycine responses is given in Fig. 2 (data from four different cells). When both cAMP and protein kinase inhibitor were included in the intracellular perfusate, glycine

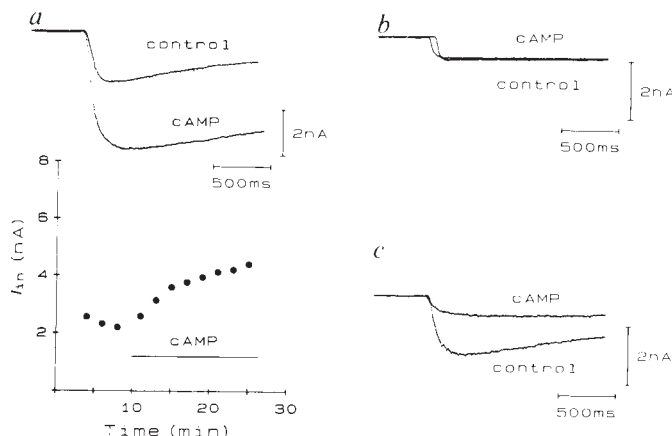


FIG. 1 Cyclic AMP-induced increase in glycine responses. *a*, Upper: Superimposed glycine-activated Cl^- currents recorded in control and cAMP-containing pipette solutions. The cell was held at -80 mV, and the inward Cl^- currents were measured as glycine was rapidly perfused onto the cell. Intracellularly applied cAMP dramatically increased glycine responses. The current was obtained 20 min after cAMP was introduced into the pipette. Lower: The changes in glycine responses with time as cAMP was introduced (indicated by the bar) to the cell. Cyclic AMP significantly increased the glycine-activated currents ($I_{\text{test}}/I_{\text{control}} = 1.70 \pm 0.13$ (s.e.); $n = 7$, $P < 0.01$). *b*, In the same cell, the response to kainate remained the same before and after cAMP application. *c*, The GABA-induced Cl^- current, in the same cell, gradually decreased with time. The currents continued decreasing when cAMP was applied to the cell.

METHODS. Neurons were isolated from the spinal trigeminal nucleus in rats using the method described before²³ with some modification. Briefly, the lower medulla and the upper cervical regions of the spinal cord were removed from the animal and cut into 500- μm transverse sections. The slices were incubated in a buffer solution which contained (mM): NaCl (120), KCl (10), glucose (25), CaCl_2 (2), MgCl_2 (6) and PIPES buffer (10), pH 7.2, at 34°C for 30 min. The tissue was transferred to an enzyme solution containing buffer solution and 2.5 mg ml^{-1} protease (type III, Sigma). After 10 min, the slices were washed with buffer solution, the spinal trigeminal nuclei were isolated from the slices and the cells were dissociated with glass pipettes. All the experiments were performed at room temperature. Whole cell currents were recorded using the patch clamp technique²⁴. The resistance of the fire-polished patch pipettes was between 2–4 M Ω . The current was sampled at 500 μs per point. The external solution contained (mM): NaCl (140), KCl (2), glucose (10), KH_2PO_4 (2), HEPES buffer (10), CaCl_2 (2) and MgCl_2 (1), pH adjusted to 7.3. The internal solution contained (mM): KF (50), KCl (70), glucose (10), HEPES (20), EGTA (10), CaCl_2 (1), ATP (5) and MgCl_2 (1), pH adjusted to 7.2. In single-channel recordings, the pipette solution was identical to the external solution except that NaCl was replaced by KCl. Thus the reversal potential for K^+ was at 0 mV. The agonists were rapidly applied externally onto the cells using the concentration clamp method previously described²⁵. To change the solution inside the pipette, a thin plastic tubing was inserted in the pipette for introducing a test solution into the cell. The tip of the pipette was filled with a small amount of control solution; the plastic tubing then was filled with the stock test solution and suction constantly applied to the pipette. During an experiment, test solution was stopped when the solution level in the pipette was doubled. Thus the final concentrations of internally applied substances were diluted to half that of the stock. By putting the end of the plastic tubing close to the tip of the pipette, the first effect of the test substance could be seen within 1–2 min of its application. The full effect usually took 5–6 min to appear. The final concentration of agonists were: glycine, $20 \mu\text{M}$; GABA, $20 \mu\text{M}$ and kainate, $250 \mu\text{M}$. The modulators used were: cAMP, $500 \mu\text{M}$; PKA, $2 \mu\text{g ml}^{-1}$; PKI, $210 \mu\text{g ml}^{-1}$; CTX, 20 nM ; GTP- γ -S, $100 \mu\text{M}$ and PTX, $1 \mu\text{g ml}^{-1}$ activated with 5 mM DTT and 2 mM NADH²⁶.

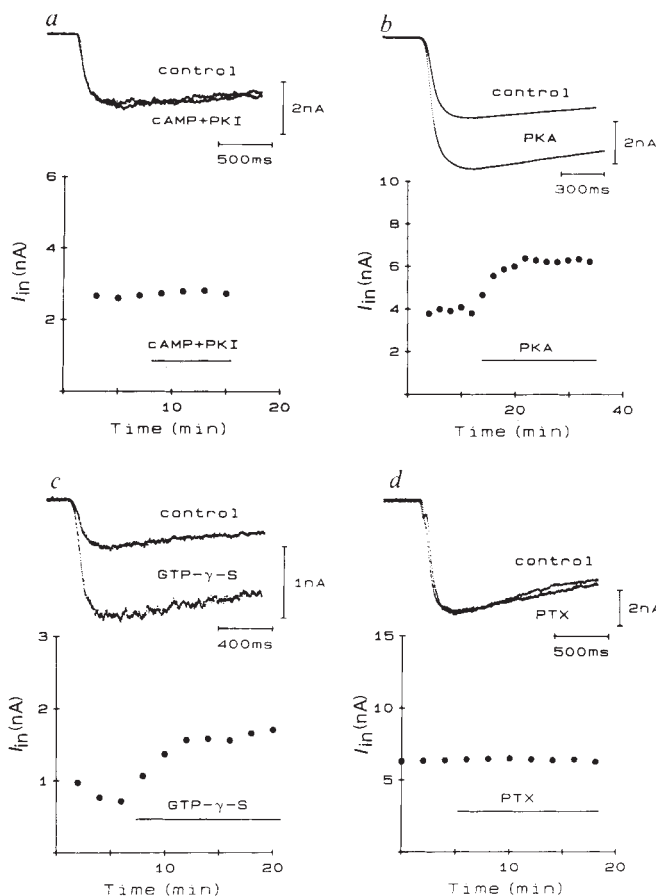


FIG. 2 Effects of intracellular cAMP + protein kinase inhibitor (PKI) (*a*), protein kinase A (PKA) (*b*), GTP- γ -S (*c*) and pertussis toxin (PTX) (*d*) on glycine-activated currents. Upper traces: the glycine-activated currents recorded before and after the chemicals were introduced into the cells. Lower curves: the time course of the changes in the peak glycine responses. Bars indicate the time period when cells were dialysed with various chemicals. Records in *a*, *b*, *c* and *d* were obtained from four different neurons. *a*, The enhancing effect of cAMP was not observed when both cAMP and protein kinase inhibitor were introduced into the cell. *b*, When PKA was introduced into the cells, glycine responses were almost doubled ($I_{\text{test}}/I_{\text{control}} = 1.95 \pm 0.19$ (s.e.); $n = 7$, $P < 0.01$). *c*, The glycine-activated currents were increased when GTP- γ -S was applied to the cells ($I_{\text{test}}/I_{\text{control}} = 2.38 \pm 0.28$ (s.e.); $n = 5$, $P < 0.01$). *d*, PTX had no effect on glycine responses.

responses stayed at the control level (Fig. 2a). When protein kinase A was introduced in the pipette solution, a large increase in glycine-induced Cl^- current was again observed. The time course of the effect induced by protein kinase A was similar to that with cAMP (Fig. 2b).

We next determined which G protein was involved in the modulation. When the non-dissociating GTP analogue GTP- γ -S was included in the pipette, the glycine-induced Cl^- current was stimulated (Fig. 2c). The glycine responses did not, however, change when pertussis toxin was introduced into the cells (Fig. 2d). When G_s protein was persistently activated by cholera toxin (CTX), the glycine response increased substantially (Fig. 3a). This increase was blocked when protein kinase inhibitor and CTX or a non-hydrolysable analogue of ATP (ATP- γ -S, 5 mM) and CTX were introduced simultaneously into the cell (data not shown). These results suggest that a G_s protein in the cAMP pathway, not G_o or G_i , is involved in modulating glycine responses.

To further confirm the specificity of cAMP action, we also studied the effect of CTX on GABA responses. Using a low concentration of GABA (10 μM) and including magnesium-ATP in the pipette, GABA responses were maintained at a stable level after an initial decrease. When they were stabilized, CTX was introduced into the pipette. In contrast to its effect on glycine responses, CTX did not increase GABA-activated currents (Fig. 3b).

To explore the mechanism of potentiation by cAMP, we determined the current-voltage relation and the dose-response of glycine in the absence or presence of cAMP. With Cl^- concentration in the extracellular solution higher than in the intracellular solution, the current-voltage relation of the glycine responses showed slight outward rectification (Fig. 4a). After cAMP treatment, the current increased and the shape of the current-voltage curve remained the same. The reversal potential of the glycine responses was also similar to that of the control (Fig. 4a). Cyclic AMP did not change the affinity of glycine to its receptor (Fig. 4b). The apparent dissociation constant for glycine was $21.8 \pm 3.8 \mu\text{M}$ ($n=5$) in the control solution and $23.8 \pm 4.7 \mu\text{M}$ ($n=5$) when cAMP was introduced into the cell.

We then determined the single-channel conductance and probability of opening by studying the effects of dibutyryl cAMP on glycine responses. In whole-cell recordings, the glycine-activated current was enhanced when neurons were incubated in 10 mM dibutyryl cAMP for 5 min (Fig. 4c). This enhancement lasted for 10–20 min even after dibutyryl cAMP was washed from the external medium (Fig. 4d). The effect of dibutyryl cAMP on glycine responses was identical to that of cAMP (Fig.

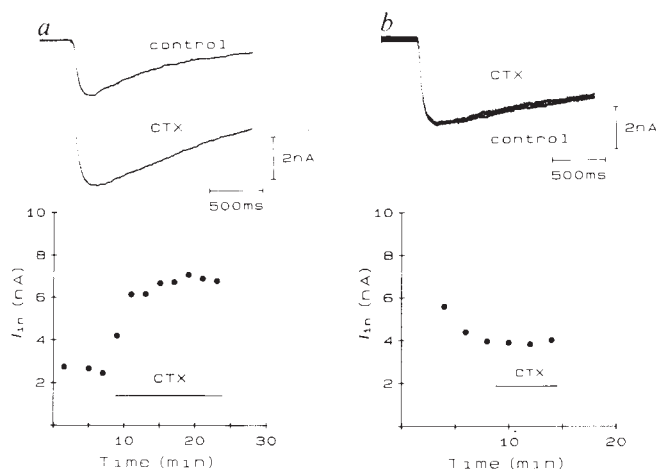


FIG. 3 Cholera toxin (CTX) enhanced the glycine responses. *a*, After introduction of CTX (20 nM) into the cells, the glycine-activated currents were enhanced ($I_{\text{test}}/I_{\text{control}} = 2.04 \pm 0.21$ (s.e.); $n=5$, $P < 0.01$). *b*, GABA responses were not altered when neurons were dialysed with CTX. The same experiments were repeated in four cells.

1a). Single glycine-activated Cl^- channel currents were recorded in cell-attached patches. With 10 μM glycine in the pipette, the inward-going channel usually had two current levels (Fig. 4e). The main level, which occurred more than 97% of the time, had an average size of $1.49 \pm 0.03 \text{ pA}$ ($n=4$) at 0 mV. The sublevel had average current of $1.06 \pm 0.03 \text{ pA}$ ($n=4$). With the resting potential around -55 mV , the estimated conductances were 27.1 pS and 19.3 pS respectively, which corresponded well with the conductances of single glycine-activated channels observed in cultured spinal neurons⁸. After putting dibutyryl cAMP in the bath, the changes in channel conductance and mean channel open time were small, but the probability of channel opening (p) increased substantially. The value of p was 0.09 ± 0.02 ($n=5$) in the absence of dibutyryl cAMP and 0.23 ± 0.06 ($n=5$) when it was present. Thus dibutyryl cAMP potentiated glycine responses by increasing the probability of opening of glycine-activated channels.

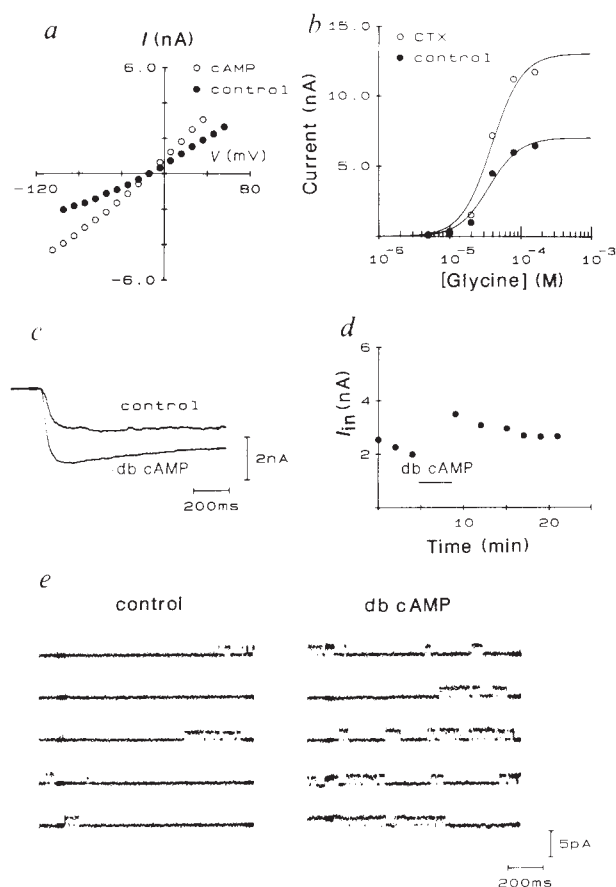


FIG. 4 Mechanism of enhancement of glycine responses by cAMP and dibutyryl cAMP (db cAMP). *a*, The current-voltage relation of glycine responses before and after cAMP treatment. The amplitude of the glycine-activated currents was increased when 500 μM cAMP was introduced into the cell. The shape of the I - V curve and the reversal potential of the glycine responses did not change. *b*, The dose-response curves for glycine in the presence or absence of intracellular CTX. The solid lines were drawn according to the equation $I = I_{\text{max}}[\text{Gly}]/(K_{\text{app}} + [\text{Gly}])$, where I_{max} is the maximal current, $[\text{Gly}]$ is the extracellular glycine concentration, K_{app} is the apparent dissociation constant. cAMP did not change the affinity for glycine. In this cell, with no cAMP present, $K_{\text{app}} = 35.6 \mu\text{M}$, $I_{\text{max}} = 7.05 \text{ nA}$, Hill coefficient = 2.0; with cAMP, $K_{\text{app}} = 39.3 \mu\text{M}$, $I_{\text{max}} = 13.07 \text{ nA}$, Hill coefficient = 2.0. *c*, Glycine-activated whole-cell currents increased after neuron was incubated in 10 mM dibutyryl cAMP. $I_{\text{test}}/I_{\text{control}} = 1.52 \pm 0.11$ (s.e.), ($n=3$, $P < 0.01$). *d*, The inward glycine-activated currents stayed elevated for 10–20 min after dibutyryl cAMP was washed away. *e*, Records of single glycine-activated channels in a cell-attached patch. Outward currents appear as upward deflections. In the control solution, the main current level was 1.6 pA, probability of opening (p) is 0.09. In solution containing dibutyryl cAMP, the current level was 1.6 pA, and $p = 0.3$.

As many hormones and transmitters act through cAMP pathways, modulation of glycine responses may prove to be a common mechanism by which the inhibitory action of this transmitter is regulated. For pain-transmission neurons, the possibility of such regulation is especially appealing. It has been established that stimulations of midbrain periaqueductal gray and rostroventral medulla inhibit the responses of spinal and medullary dorsal horn neurons to noxious stimuli^{3,4,15}. The mechanism of this inhibition is unclear. As serotonin (5-HT)

and noradrenaline are present in the descending pathways¹⁶⁻²⁰ and inhibit glutamate- and pinch-induced activities in spinothalamic tract neurons⁶, they seem to have important roles in the control of pain. Both α_2 -adrenergic and 5-HT₁ receptors are thought to be involved^{17,18,21}. Noradrenaline and 5-HT are known to change the cAMP level in the target cells²². These transmitters could change the cellular responses to glycine through protein phosphorylation, which in turn would affect pain transmission at medulla and spinal levels. □

Received 7 August; accepted 12 September 1990.

1. Curtis, D. R., Hosli, L. & Johnston, G. A. R. *Nature* **215**, 1502-1503 (1967).
2. Curtis, D. R., Duggan, A. W. & Johnston, G. A. R. *Brain Res.* **14**, 759-762 (1967).
3. Sessle, B. J., Hu, J. W., Dubner, R. & Lucier, G. E. *J. Neurophysiol.* **45**, 193-207 (1981).
4. Dubner, R. & Bennett, G. J. A. *Rev. Neurosci.* **6**, 381-418 (1983).
5. Hu, J. W. & Sessle, B. J. *J. Neurophysiol.* **52**, 39-53 (1984).
6. Willcockson, W. S., Chung, J. M., Hori, Y., Lee, K. H. & Willis, W. D. *J. Neurosci.* **4**, 732-740 (1984).
7. Barker, J. L. & Nicoll, R. A. *J. Physiol., Lond.* **228**, 259-277 (1973).
8. Bormann, J., Hamill, O. P. & Sakmann, B. *J. Physiol., Lond.* **385**, 243-286 (1987).
9. Pfeiffer, F., Graham, D. & Betz, H. *J. Biol. Chem.* **257**, 9389-9393 (1982).
10. Graham, D., Pfeiffer, F. & Betz, H. *Biol. Chem.* **264**, 990-994 (1985).
11. Betz, H. & Becker, C.-M. *Neurochem. Int.* **13**, 137-146 (1988).
12. Grenningloh, G. *et al.* *Nature* **328**, 215-220 (1987).
13. Grenningloh, G. *et al.* *Nature* **330**, 25-26 (1987).
14. Stelzer, A., Kay, A. R. & Wong, R. S. *Science* **241**, 339-341 (1988).
15. Basbaum, A. I. & Fields, H. L. A. *Rev. Neurosci.* **7**, 309-338 (1984).

16. Westlund, K. N. & Coulter, J. D. *Brain Res. Rev.* **2**, 235-264 (1980).
17. Yash, T. L. & Reddy, S. V. R. *Anesthesiology* **54**, 451-467 (1981).
18. Yezierski, R. P., Wilcox, T. K. & Willis, W. D. *J. Pharmac. exp. Ther.* **220**, 266-277 (1982).
19. Westlund, K. N., Bowker, R. W., Ziegler, M. G. & Coulter, J. D. *Brain Res.* **292**, 1-16 (1984).
20. Duggan, A. W. *Phil. Trans R. Soc. B* **308**, 375-391 (1985).
21. Fleetwood-Walker, S., Mitchell, R., Hope, P., Molony, V. & Iggo, A. *Brain Res.* **334**, 243-254 (1985).
22. McGeer, P. L., Eccles, J. C. & McGeer, E. G. in *Molecular Neurobiology of the Mammalian Brain* 149-173 (Plenum, New York, 1987).
23. Huang, L.-Y. M. *J. Physiol., Lond.* **411**, 161-177 (1989).
24. Hamill, O., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
25. O'Dell, T. J. & Christensen, B. N. *J. Neurophysiol.* **61**, 162-172 (1989).
26. Kurachi, Y., Nakajima, T. & Sugimoto, T. *Pflügers Arch. ges. Physiol.* **407**, 264-274 (1986).

ACKNOWLEDGEMENTS. This work is supported by grants from the NIH (L.M.H.). We thank L. Chen for comments, S. Y. Wong and K. Forni for technical assistance and L. Simmons and M. Watson for preparing the manuscript.

Circulating CD8⁺ cytotoxic T lymphocytes specific for HTLV-I pX in patients with HTLV-I associated neurological disease

Steven Jacobson*, Hisatoshi Shida†, Dale E. McFarlin*
Anthony S. Fauci† & Scott Koenig‡

*Neuroimmunology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892, USA

†Institute of Virology, Kyoto University, Kyoto 606, Japan

‡Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

THE human T-lymphotropic virus type I (HTLV-I), the first human retrovirus to be characterized, is associated with adult T-cell leukaemia and a chronic progressive disease of the central nervous system termed tropical spastic paraparesis, or HTLV-I-associated myelopathy¹⁻⁴. Only 1% of individuals infected with HTLV-I develop clinical disease however⁵. The various manifestations of an HTLV-I infection may be related to differences in the genetic backgrounds of individuals⁶, infection with variant strains of HTLV-I (ref. 7), differences in viral tropism⁸ or host immune response to the virus. Whereas the humoral response to HTLV-I is well characterized^{3,4,9,10}, little is known about the human cellular immune response, such as the production of cytotoxic T lymphocytes¹¹⁻¹³. Here we report the presence of high levels of circulating HTLV-I-specific cytotoxic T lymphocytes in patients with HTLV-I associated neurological disease but not in HTLV-I seropositive individuals without neurological involvement. These cytotoxic T lymphocytes are CD8⁺, HLA class I-restricted and predominantly recognize the HTLV-I gene products encoded in the regulatory region pX. These findings suggest that HTLV-I-specific cytotoxic T lymphocytes may contribute to the pathogenesis of associated neurological disorders associated with HTLV-I.

Peripheral blood lymphocytes (PBL) were isolated from HTLV-I seropositive individuals (Table 1) and tested for their capacity to lyse autologous Epstein-Barr virus-transformed lym-

phoblastoid cell lines (LCL) infected with HTLV-I recombinant vaccinia virus expressing various HTLV-I gene products¹⁴⁻¹⁷ (Fig. 1). Circulating lymphocytes from patient R.S., suffering from tropical spastic paraparesis, lysed autologous LCL infected with either the HTLV-I p27x or p40x, but not the p21x construct (Fig. 1a). Lower levels of lysis were also observed against LCL expressing the HTLV-I env gene products. As both the p27x and p40x constructs express the HTLV-I tax protein¹⁴⁻¹⁷ (Fig. 1a legend), this suggests that the HTLV-I tax protein (p40x) is the predominant HTLV-I gene product recognized, although reactivity with the rex protein (p27x) cannot be excluded.

The demonstration of HLA restriction¹⁸ is also evidence for specific CTL recognition. PBL from patient R.S. were used as effector lymphocytes on a panel of HLA-typed allogeneic LCL. Lysis was observed with either the autologous LCL or an allogeneic LCL that shared only the HLA class I allele A2 (Fig. 1b). To confirm that this response was HLA class I-restricted, a transfected LCL target that only expressed HLA-A2.1 (ref. 19) was used. After infection with either the HTLV-I p27x or p40x vaccinia recombinant, the target cells transfected with HLA-A2.1 were susceptible to lysis with PBL from R.S., whereas a similarly infected neomycin-transfected LCL was not (Fig. 1c and d).

To determine the T-cell phenotype of circulating HTLV-I-specific CTL, fresh PBL of patient R.S. were sorted into CD8⁺ and CD4⁺ populations and directly used as effectors on autologous LCL infected with the HTLV-I p40x vaccinia recombinant. Although unfractionated lymphocytes from R.S. showed detectable levels of cytotoxicity, this lysis was greatly enhanced in the purified CD8⁺ population (Fig. 1e). Purified CD4⁺ cells did not lyse these targets.

We looked for circulating HTLV-I-specific CTL in other individuals seropositive for HTLV-I (Table 1). Tropical spastic paraparesis patient R.B. and the narcoleptic patient S.T. had detectable levels of circulating CTL that recognized autologous LCL infected with either the HTLV-I p27x or p40x vaccinia constructs (Fig. 2a and b). In contrast to patients with neurological disease, fresh PBL from seropositive, asymptomatic donors and patients without neurological disease (Table 1) did not lyse autologous LCL infected with any of the HTLV-I-vaccinia constructs (Fig. 2c and d). HTLV-I-specific CTL activity was not found in controls seronegative for HTLV-I (data not shown).

TABLE 1 HTLV-I-specific CTL responses

Donor Age/sex	Clinical status*	CTL response							
		env	Fresh			CD8 ⁺ line			
			gag	p40x	p21x	env	gag	p40x	p21x
R.S. 45/M	TSP	+	—	++	—	+	—	++	—
R.B. 60/M	TSP	—	—	++	—	ND	ND	ND	ND
E.G. 45/F	TSP	ND	ND	ND	ND	—	—	++	+
C.S. 34/M	TSP	ND	ND	ND	ND	+	+	++	—
S.T. 38/F	Narcolepsy	+	—	++	—	+	—	++	—
K.S. 43/F	Asymptomatic	—	—	—	—	—	ND	—	—
J.W. 43/F	Asymptomatic	—	—	—	—	—	—	—	—
C.H. 33/F	Asymptomatic	—	—	—	—	—	—	—	—
J.B. 32/F	Adult T-cell leukaemia	—	—	—	—	—	ND	—	—
P.P. 48/M	Strongyloidiasis	—	—	—	—	—	ND	—	—

++, Major HTLV-I proteins recognized; +, detectable response (>10%) to other HTLV-I proteins; —, no response above background; ND, not done.

* TSP, patients fulfilled clinical criteria for diagnosis of tropical spastic paraparesis/HTLV-I associated myelopathy^{4,13} and have cerebrospinal fluid findings consistent with this disorder. Patient S.T. has clinical features and encephalographic findings consistent with narcolepsy; neurological examination showed brain stem abnormalities and normal cerebrospinal fluid.

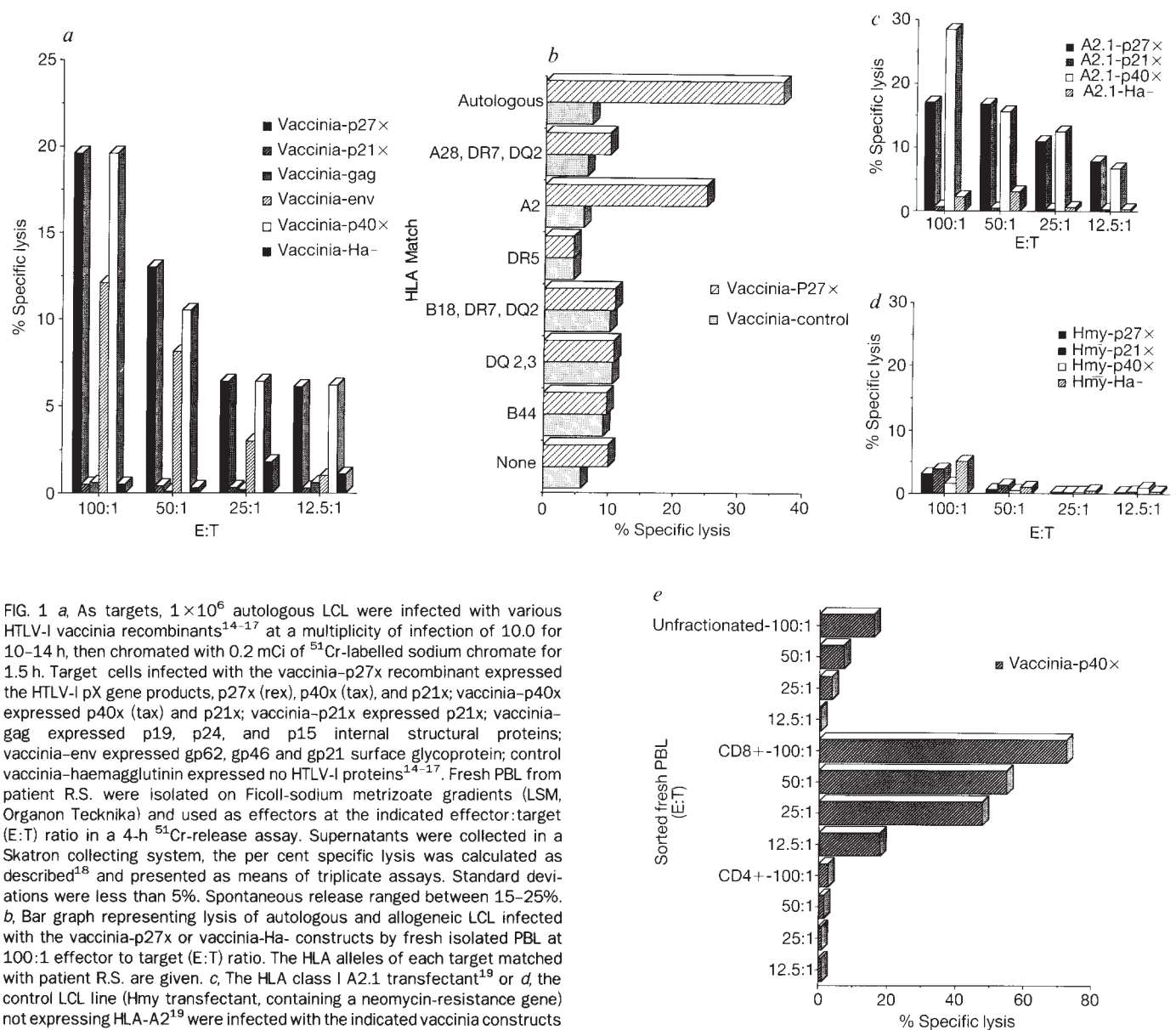


FIG. 1 *a*, As targets, 1×10^6 autologous LCL were infected with various HTLV-I vaccinia recombinants¹⁴⁻¹⁷ at a multiplicity of infection of 10.0 for 10–14 h, then chromated with 0.2 mCi of ^{51}Cr -labelled sodium chromate for 1.5 h. Target cells infected with the vaccinia-p27x recombinant expressed the HTLV-I pX gene products, p27x (rex), p40x (tax), and p21x; vaccinia-p40x expressed p40x (tax) and p21x; vaccinia-p21x expressed p21x; vaccinia-gag expressed p19, p24, and p15 internal structural proteins; vaccinia-env expressed gp62, gp46 and gp21 surface glycoprotein; control vaccinia-haemagglutinin expressed no HTLV-I proteins¹⁴⁻¹⁷. Fresh PBL from patient R.S. were isolated on Ficoll-sodium metrizoate gradients (LSM, Organon Technika) and used as effectors at the indicated effector:target (E:T) ratio in a 4-h ^{51}Cr -release assay. Supernatants were collected in a Skatron collecting system, the per cent specific lysis was calculated as described¹⁸ and presented as means of triplicate assays. Standard deviations were less than 5%. Spontaneous release ranged between 15–25%. *b*, Bar graph representing lysis of autologous and allogeneic LCL infected with the vaccinia-p27x or vaccinia-Ha- constructs by fresh isolated PBL at 100:1 effector to target (E:T) ratio. The HLA alleles of each target matched with patient R.S. are given. *c*, The HLA class I A2.1 transfectant¹⁹ or *d*, the control LCL line (Hmy transfectant, containing a neomycin-resistance gene) not expressing HLA-A2¹⁹ were infected with the indicated vaccinia constructs and used as targets. Effectors were fresh isolated PBL from patient R.S. Spontaneous release of targets ranged between 10–20%. *e*, Fresh isolated PBL from patient R.S. were stained with monoclonal antibodies Leu3(a and b) and Leu2 (Becton-Dickinson) and positively sorted on an Epics flow cytometer (Coulter). Resultant populations were greater than 98% pure.

Unfractionated, CD8⁺, and CD4⁺ populations were used as effectors at the indicated E:T ratios. Targets were autologous LCL infected with the vaccinia-p40x construct. Target cells infected with the control vaccinia-haemagglutinin recombinant were not lysed (data not shown).

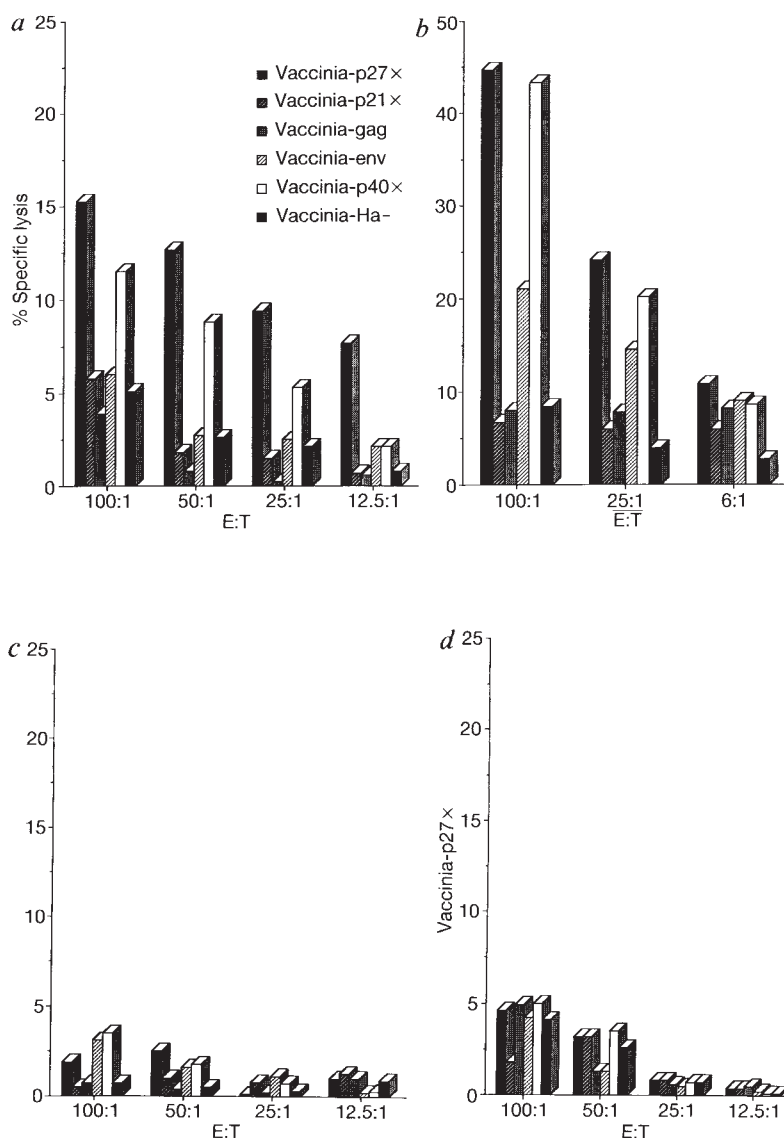


FIG. 2 Autologous LCL were infected with the indicated HTLV-I vaccinia recombinants and used as targets. As effectors, fresh isolated PBL from HTLV-I seropositive individuals were used at the indicated E:T ratios. a, Patient R.B.; b, narcoleptic patient S.T.; c, asymptomatic donor K.S. (wife of patient R.S.); and d, asymptomatic donor J.W. All experiments were repeated two or three times.

Circulating CTL have been demonstrated in other retroviral systems²⁰ but in those experiments, as well as in our present study, the lytic ability of the CTL was greatly reduced if cryopreserved PBL were used. To circumvent the need to use fresh PBL, thawed lymphocytes were sorted into CD8⁺ and CD4⁺ T-cell populations and nonspecifically expanded with phytohaemagglutinin in interleukin-2. The CD8⁺ lines from patients R.S. and S.T. had the same pattern of lysis as fresh PBL on autologous targets infected with the same HTLV-I-vaccinia recombinants (Table 1). No lysis was observed with the expanded CD4⁺ population. Cryopreserved lymphocytes from two additional tropical spastic paraparesis patients, E.G. and C.S., were sorted similarly (Fig. 3). CD8⁺ T-cell lines from both E.G. and C.S. lysed targets infected with either the HTLV-I p40x or p27x construct (Fig. 3a and c). A lesser degree of lysis was also observed on targets expressing HTLV-I env or gag (patient C.S.) or the p21x protein (patient E.G.). The CD4⁺ T-cell lines from E.G. and C.S. were not cytolytic against targets expressing any HTLV-I protein (Fig. 3b and d). CD8⁺ and CD4⁺ lines from two additional HTLV-I seropositive, asymptomatic donors (J.W. and C.H.) also showed no CTL activity (Table 1).

Transient circulating CTL have been detected in a number of acute viral infections²¹. The effector cells were CD8⁺, HLA-class I-restricted and represent one of the earliest cellular immune

responses against viral infection. In convalescents, however, the demonstration of circulating virus-specific CTL¹⁸ requires *in vitro* antigenic stimulation and expansion. Only one other chronic viral disease (AIDS) has been associated with virus-specific circulating CTL that are demonstrable without *in vitro* priming²⁰. The effector CTL were predominantly CD8⁺ and recognized a variety of human immunodeficiency virus antigens.

Although HTLV-I-specific CTL have been cultured from patients with adult T-cell leukaemia¹¹⁻¹² the HTLV-I proteins recognized in these studies were not identified. Moreover, fresh, circulating CTL were not noted. In the present study, the inability to demonstrate circulating CD8⁺, HTLV-I-specific CTL from one patient with adult T-cell leukaemia was not surprising as this individual had T-cell clonal expansion of CD4⁺ leukaemic cells characteristic of this disease²².

The presence of such high levels of circulating CTL in patients with neurological disease raises the possibility that this cellular response may be involved in the pathogenesis of tropical spastic paraparesis. Alternatively, the high level of CTL activity may be related to an increased viral load in clinically symptomatic patients²³. HTLV-I seropositive, asymptomatic individuals with a high viral load might, similarly, have detectable levels of HTLV-I-specific CTL, although this was not observed in the limited number of healthy individuals studied. If HTLV-I-specific CTL actually have a role in the pathogenesis of the

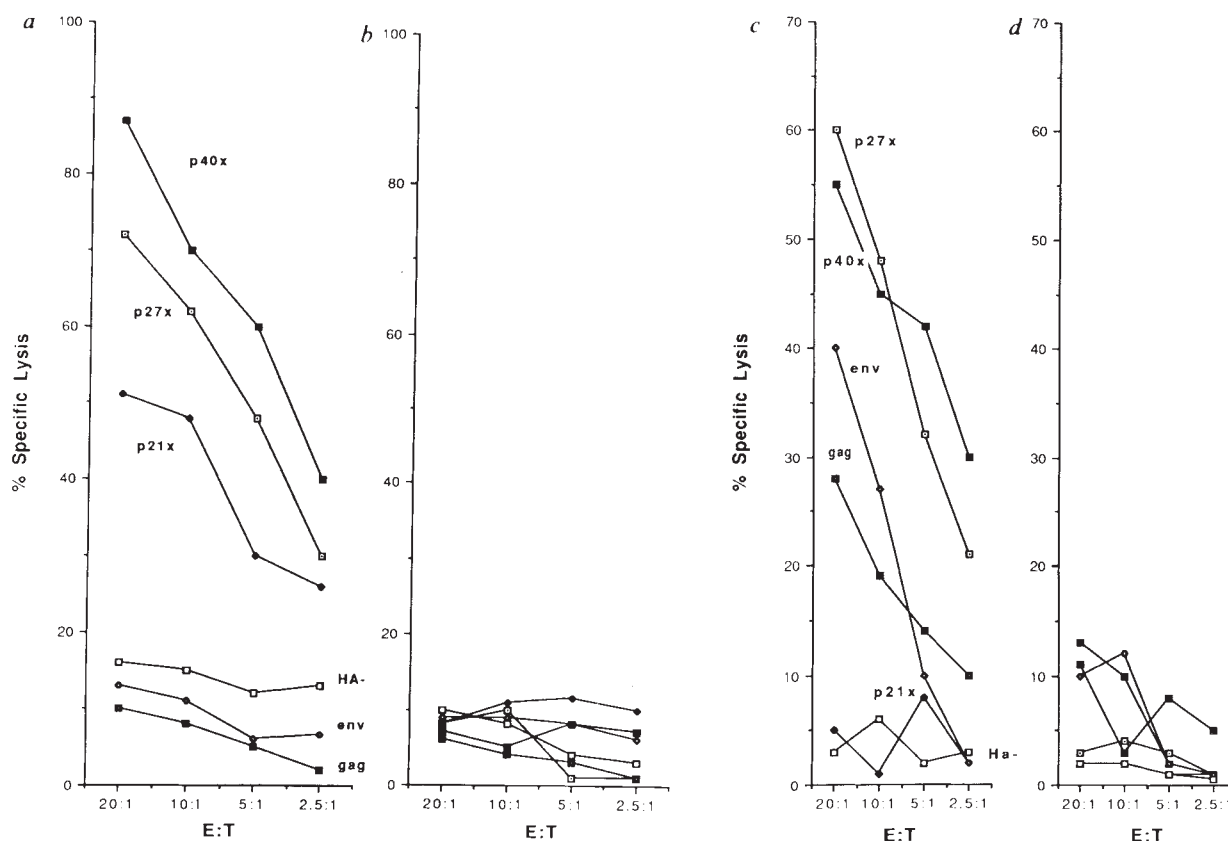


FIG. 3 Cryopreserved PBL were thawed and stained with monoclonal antibodies Leu3(a and b) and Leu2a and sorted into purified CD8⁺ and CD4⁺ populations. Sorted populations were greater than 98% pure for each T-cell surface marker. Cells were expanded in culture by plating 1×10^6 sorted T-cells with 5×10^5 autologous irradiated PBL (4,000 R) in RPMI-1640 plus 10% fetal calf serum in 24-well Costar plates supplemented with $1.0 \mu\text{g ml}^{-1}$ phytohaemagglutinin, 100 units ml^{-1} recombinant interleukin-2

(Cetus) and 10% purified IL-2 (Electronucleonics). Culture medium (RPMI-1640, 10% FCS, and IL-2) was changed three times a week. After 2–4 weeks in culture, expanded populations were used as effectors on autologous LCL infected with the indicated HTLV-I vaccinia recombinants. a, Purified CD8⁺ line and b, CD4⁺ line from patient E.G. c, Purified CD8⁺ line and d, CD4⁺ line from patient C.S.

neurological condition, then these cells would be expected to be present in lesions. CD8⁺ T-cells have been observed in central nervous system lesions of one patient with tropical spastic paraparesis⁸.

Cellular peptide composition governed by major histocompatibility complex class I molecules

Kirsten Falk, Olaf Röttschke
& Hans-Georg Rammensee*

Abteilung Immungenetik, Max-Planck-Institut für Biologie,
Corrensstrasse 42, D-7400 Tübingen, Germany

MAJOR histocompatibility complex (MHC) class I molecules present peptides derived from cellular proteins to cytotoxic T lymphocytes (CTLs)^{1–5}, which check these peptides for abnormal features. How such peptides arise in the cell is not known. Here we show that the MHC molecules themselves are substantially involved in determining which peptides occur intracellularly: normal mouse spleen cells identical at all genes but MHC class I express different patterns of peptides derived from cellular non-MHC proteins. We suggest several models to explain this influence of MHC class I molecules on cellular peptide composition.

* To whom correspondence should be addressed.

Received June 19; accepted 10 September 1990.

- Poelisz, B. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7415–7419 (1980).
- Wong-Staal, F. & Gallo, R. C. *Nature* **317**, 395–413 (1985).
- Gessain, A. *et al. Lancet* **ii**, 407–410 (1985).
- Osame, M. *et al. Lancet* **i**, 1031–1032 (1986).
- Murphy, E. L. *et al. Int. J. Cancer* **43**, 250–253 (1989).
- Usuku, K. *et al. Ann. Neurol. (Suppl.)* **23**, S143–S150 (1988).
- Jacobson, S., Raine, C. S., Mingioli, E. S. & McFarlin, D. E. *Nature* **331**, 540–542 (1988).
- Moore, W. G. R., Traugott, U., Schienberg, L. C. & Raine, C. S. *Neurology* **26**, 523–530 (1989).
- Palker, T. J. *et al. J. Immun.* **142**, 971–978 (1989).
- Kurata, A. *et al. J. Immun.* **143**, 2024–2030 (1989).
- Kannagi, M. K. *et al. J. Immun.* **130**, 2942–2946 (1983).
- Mitsuya, H. L. *et al. J. exp. Med.* **158**, 994–999 (1983).
- Jacobson, S., Gupta, A., Mattson, D., Mingioli, E. S. & McFarlin, D. E. *Ann. Neurol.* **27**, 149–155 (1990).
- Shida, H. *et al. EMBO J.* **6**, 3379–3384 (1987).
- Nam, S. H., Kidokoro, M., Shida, H. & Hatanaka, M. *J. Virol.* **62**, 3718–3728 (1988).
- Siomi, H. *et al. Cell* **55**, 197–209 (1988).
- Tanaka, Y., Tozawa, H., Koyanagi, Y. & Shida, H. *J. Immun.* **144**, 4202–4211 (1990).
- Townsend, A. & Bodmer, H. A. *Rev. Immun.* **7**, 601–624 (1989).
- Shimojo, N. *et al. Intern. Immun.* **2**, 193–200 (1990).
- Plata, F. *et al. Immunodeficiency Rev.* **1**, 227–246 (1989).
- Fleischer, B. & Kreth, H. W. *Hum. Immun.* **7**, 239–248 (1983).
- Waldmann, T. A. *et al. J. clin. Invest.* **73**, 1711–1718 (1983).
- Gessain, A. *et al. Blood* **75**, 428–433 (1990).

ACKNOWLEDGEMENTS. We thank S. Houff, T. Waldmann, H. Emsellem and F. Neva for providing patients; G. Jones, Y. A. Brewah, A. J. Newell, E. Boone and J. S. Reuben for technical assistance; M. Baseler and C. Brudzyn (PRI Inc.) for help with cell sorting; and B. Biddison for transfected cell lines.

Normal somatic cells simultaneously present many peptides derived from normal cellular proteins on their cell surface MHC class I molecules. Such peptides are detectable with CTLs derived from an MHC-matched individual that can recognize a given peptide as foreign because of a polymorphism or of a difference in expression of that peptide^{3,4,6}. Such peptides are, for historical reasons, called minor histocompatibility antigens. We used C57BL/6 (H-2^b) derived, H-2^b-restricted CTLs specific for three different mouse minor histocompatibility antigens (H-4^b, H-Y, and an unmapped BALB.B antigen called mapki, see Table 1) to detect several members of the pool of naturally occurring peptides in normal cells.

Peptides were extracted from spleen cells of C57BL/6 (H-4^a) females, of BALB.B (H-2^b, H-4^b-like) males and of BALB/c (H-2^d) males by acid elution⁴. Note that entire cells were used for extraction, and no attempt was made to enrich for MHC-molecules. Extracted peptides were separated by HPLC and tested for CTL recognition. BALB.B male cells contained all three peptides (Fig. 1a), whereas C57BL/6 female cells did not contain peptides recognized by any of our three CTL lines (Fig.

1c). All peak fractions from BALB.B male cells were recognized in a concentration-dependent fashion by the respective CTL (Fig. 1d-g). BALB/c (H-2^d) males, which express the same minor histocompatibility genes as BALB.B males, contained, to our surprise, none of these peptides (Fig. 1b), apart from one exception. A small, but reproducible peak (see below) around fraction 26 of both BALB.B and BALB/c peptide extracts was recognized by the H-4^b-specific line B21W9.

To test further the influence of MHC genes on cellular peptide composition, total peptides were extracted from MHC-recombinant BALB.5R (K^bD^d) and BALB.HTG (K^dD^b) male spleen cells. K^b-expressing BALB.5R males contained the K^b-restricted H-4^b and mapki peptides, but not the D^b-restricted male specific peptide (Fig. 1h). By contrast, D^b-expressing BALB.HTG males contained the male peptide, but not the mapki peptide, and not the main H-4^b peptide (Fig. 1i). These experiments indicated an influence of genes closely linked to H-2 class I genes on cellular peptide composition. The K^b-restricted H-4^b main peptide was also detected in (BALB/c × C57BL/6)F₁, but not in (BALB/c × B6.C-H-2^{bml})F₁ extracts, although the

FIG. 1 Occurrence of minor histocompatibility peptides in cells is MHC class I dependent. Acid extracts prepared from spleen cells of BALB.B males (a), BALB/c males (b) or C57BL/6 females (c) were separated by reversed-phase HPLC. Individual fractions were incubated with EL 4 (H-2^b) target cells and tested for recognition by H-Y-specific, D^b-restricted CTL 11P9 (filled triangles); by H-4^b-specific, K^b-restricted CTL B21W9 (filled circles); by the BALB.B-specific, K^b-restricted CTL clone mapki (filled diamonds); and without CTL (open circles). Positive fractions were retested in titrated amounts for recognition by the respective CTLs: fraction 26 with B21W9 (d), fraction 27 with mapki (e), fraction 28 with 11P9 (f) and fraction 29 with B21W9 (g). Further acid extracts prepared from male spleen cells from the following mouse strains were separated and analysed as above. h, BALB.5R; i, BALB.HTG; j, (BALB/c × C57BL/6)F₁; k, (BALB/c × B6.C-H-2^{bml})F₁; l, (C57BL/6 × DBA/2)F₁; m, DBA/2. For genotypes of these strains, see Table 1. In some experiments, clones B30X9-4a/34 (anti H-4^b) (filled squares) and 11P9-4/01 (anti H-Y) (filled, inverted triangles) were used. Full lines, specific ⁵¹Cr release (%); dashed lines, absorbance at 220 nm. METHODS. Two to four spleens each of the respective strain were homogenized in 0.1% trifluoroacetic acid (TFA) as described⁴. Soluble material of M_r < 5,000 (collected by gel filtration) was separated on a reversed-phase HPLC column (Suprapac pepS, Pharmacia LKB). Eluent A, 0.1% TFA; eluent B, acetonitrile containing 0.1% TFA. Gradient, 0–60%B; flow rate, 1 ml min⁻¹; fraction size, 1 ml. Fractions were dried, suspended in 0.5 ml phosphate buffered saline (PBS), incubated with ⁵¹Cr-labelled EL 4 cells and tested for recognition by the respective CTL in a standard ⁵¹Cr-release assay. Briefly, 10⁴ EL 4 cells preincubated for 90 min with the respective fraction (a–c and h–m, undiluted, d–g, diluted) were incubated for 6 h (37 °C, 5% CO₂) with 4 × 10⁴–5 × 10⁵ effector CTL in a total volume of 200 µl α-MEM medium containing 10% fetal calf serum. Radioactivity released into the supernatant was determined in a γ-counter. Spontaneous ⁵¹Cr release of EL 4 target cells ranged from 15.1–25.3%. For information on CTL lines and clones, see Table 1.

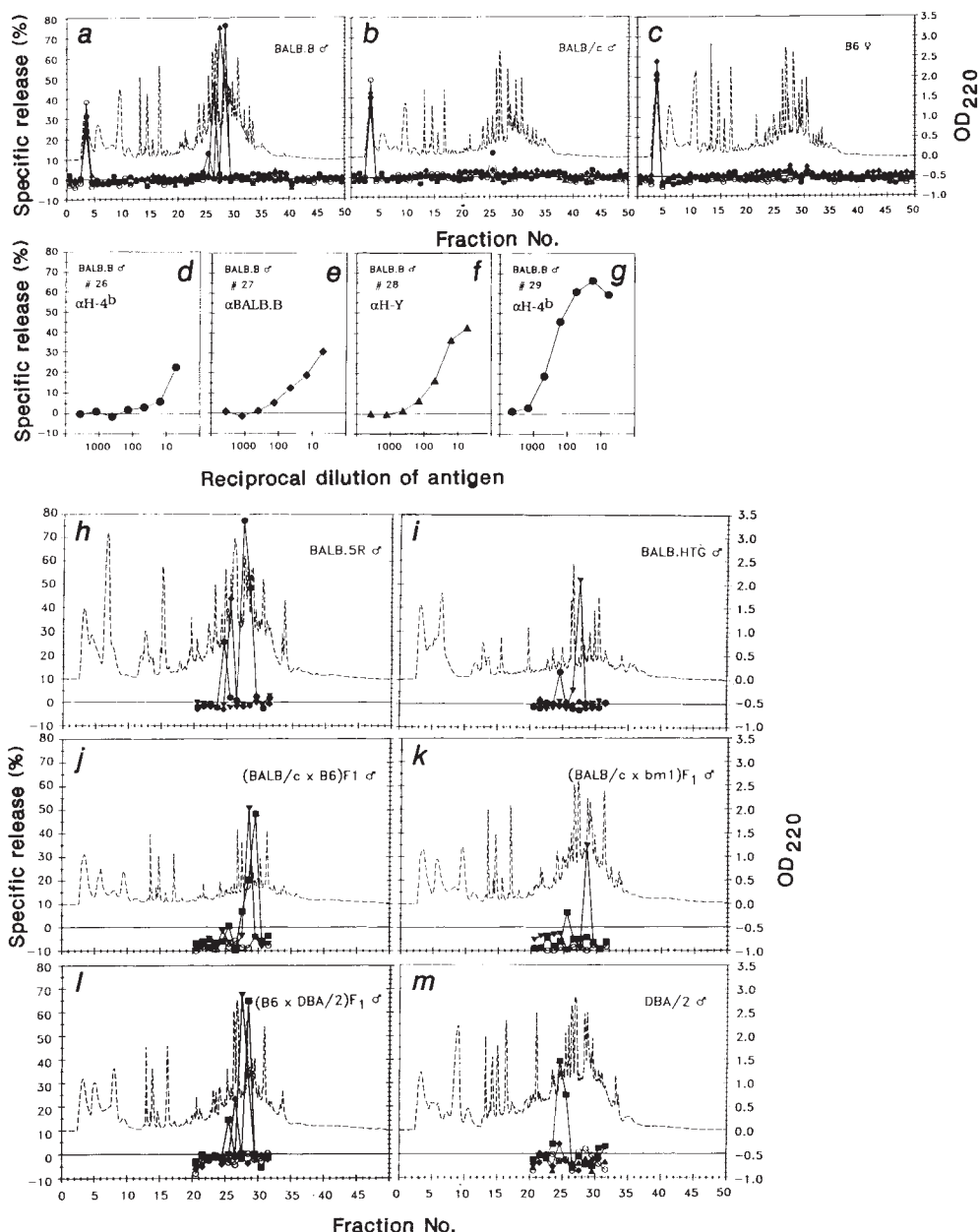


TABLE 1 Influence of MHC class I genes on peptides contained in normal cells

Mouse		H-2 ^k	H-4 ^d	Recognition of cells by CTL*					mapki αBALB/B clone	fr.29 H-4 ^b main	Peptides eluted from cells†		
				B21W9 αH-4 ^b line	B30X9-4a/34 αH-4 ^b clone	11P9 αH-Y line	11P9-4/01 αH-Y clone				fr.26 H-4 ^b pre	fr.28 H-Y	fr.27 mapki
C57BL/6	♀	bbbb	a	—	—	—	—	—	—	0.2/0.8(2)	—2.0/—0.6(2)	1.2	5.3
C57BL/6	♂	bbbb	a	—	—	+	+	—	—	0/0.7(2)	0.7/0.7(2)	64.6	ND
BALB.B	♀	bbbb	§	+	+	—	ND	+	+	67.1	26.9	0.9/1.1(2)	ND
BALB.B	♂	bbbb	§	+	+	+	+	+	+	53.1/86.4(6)	11.3/28.5(6)	30.8/75.2(6)	24.5/46.9(3)
BALB/c	♂	dddd	§	—	—	—	—	—	—	0.7/3.9(3)	12.1/32.4(3)	1.4/1.7(2)	—2.0/—2.1(2)
BALB.5R	♂	bb ^b / ₁ d	§	+	+	—	—	+	+	77	25.4	—2.2	43.6
BALB.HTG	♂	dddd	§	—	—	+	+	—	—	0.1	13.1	74.5	—0.4
129/Sv	♂	bbbb	b	+	+	+	+	+	+	62.9/91.9(3)	—2.8/11.6(3)	49.5/55.7(2)	32.6
B10.129-H-4 ^b	♂	bbbb	b	+	+	+	+	+	+	77.0/83.9(3)	22.1/36.4(2)	68.9	1.9
B6.C-H-2 ^{bm1}	♂	bm1bbb	a	—	—	+	+	—	—	ND	ND	ND	ND
(BALB/c × B6.C-H-2 ^{bm1}) _F ₁	♂	bm1bbb	b	—	—	—	—	—	—	—6.5/—1.5	0.9/6.3	29.1/34.8	—6.3/2.9
(BALB/c × C57BL/6) _F ₁	♂	bbbb	b	+	+	+	+	—¶	—	48.3	0.6	50.8	—0.8¶
DBA/2	♂	dddd	§	—	—	—	—	—	—	—4.6	39.2	—4.1	—4.1
DBA/1	♂	qqqq	§	—	—	—	—	—	—	—4.4	24.2	—2.4	—2.6
(C57BL/6 × DBA/2) _F ₁	♂	bbbb	a	+	+	+	+	—¶	—	60.5/64.8(2)	14.4/17.3(2)	59.7/67.4(2)	14.2/23.1(2)¶
CBA/LacSto	♂	kkkk	a	—	—	—	—	—	—	—2.1	—1.5	—1.5	—4.0

* Lysis of concanavalin A induced blast cells of the respective strain. These data are compiled from several experiments, part of which have been published^{3,4}. The CTL line B21W9 (ref. 3) is K^b-restricted and specific for a chromosome 7-dependent H-4^b peptide expressed in BALB.B mice⁴. 11P9 is D^b-restricted and specific for a Y-chromosome-dependent male specific peptide⁴. The CTL clone mapki is K^b-restricted and recognizes a BALB.B encoded peptide, whose gene has not been mapped⁵. The H-4^b and H-Y specific clones were derived by limiting dilution at 1 cell/well from the line B30X9 (ref. 3) or from 11P9.

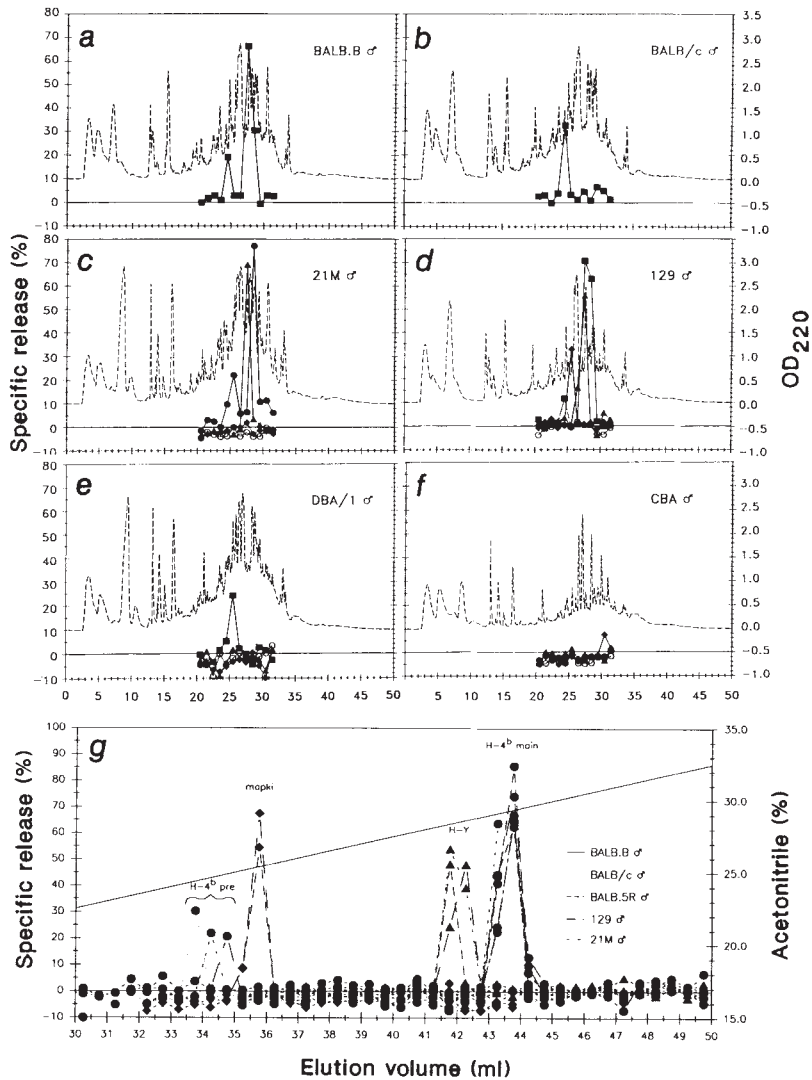
† Lysis of EL 4 target cells incubated with CTL and the respective HPLC-fraction from acid extracts. Indicated is the lowest/highest lysis observed with different preparations and the number (in brackets) of preparations.

‡ Alleles at H-2K, A, E or D/L genes, or at H-4 (refs 7, 17).

§ Not a or b. H-4^b-specific CTL, however, recognize BALB.B (ref. 3) and (C57BL/6 × DBA/2)_F₁ cells (this table).

¶ The clone mapki is relatively inefficient in lysing con A-induced blast cells, although peptide-incubated target cells are efficiently lysed³. Lysis of heterozygous cells could not be detected at all, although the antigen is at least in (B6 × DBA/2)_F₁ cells, as shown by the respective peptide eluted.

FIG. 2 Further analysis of MHC-independent and MHC-dependent peptides. a–f, Acid extracts prepared from spleen cells of males from different mouse strains were separated as in Fig. 1. Full line, specific ⁵¹Cr release (%); dashed, absorbance at 220 nm. a, BALB.B; b, BALB/c; c, B10.129-H-4^b (21M); d, 129/Sv; e, DBA/1; f, CBA/LacSto; g, fractions containing the H-4^b pre- and main peptide, the H-Y, or the mapki peptide collected from acid extracts of BALB.B (long dashed line); BALB/c (dotted line), BALB.5R (short dashed line); 129/Sv (long/short dashed line) or B10.129-H-4^b (21M) (short dashed line with double dots) male spleen cells were rechromatographed on a reversed-phase HPLC column under conditions giving higher resolution (acetonitrile gradient, solid line) and tested for CTL recognition. CTLs used for detection of peptides were the H-4^b-specific clone B30X9-4a/34 (filled squares) or B21W9 (filled circles), 11P9 (anti H-Y) (filled triangles) or mapki (anti BALB.B) (filled diamonds). In g, eluents A and B are as in Fig. 1. Gradient, 0–5 min, 0% B; 5–15 min, linear increase to 15% B; 15–60 min, increase to 37.5% B. Flow rate, 1 ml min⁻¹; fraction size, 0.5 ml.



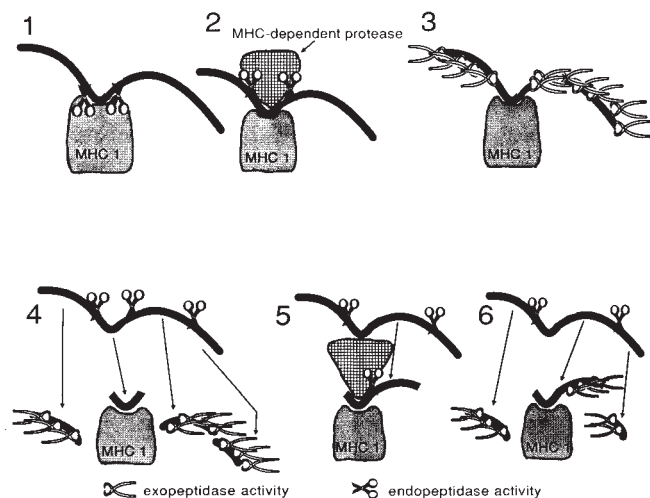


FIG. 3 Six models for the involvement of MHC class I molecules in intracellular protein degradation. Each cartoon shows an MHC class I molecule, an unfolded polypeptide chain and sites of exopeptidase or endopeptidase activities. Cellular compartmentalization, which adds additional complexity to the matter, is not considered. Model 1 depicts an MHC molecule as a protease able to cut, within one polypeptide chain, at two sites simultaneously. The model assumes binding of unfolded polypeptide chains to the peptide binding groove of MHC class I. Model 2 assumes the MHC molecule to serve as a template for an MHC-dependent protease. The cutting sites are the same as in model 1. Model 3 is similar to model 1, with the difference that no specific cutting sites are required. The polypeptide parts hanging out of the binding groove are trimmed by exopeptidases. In this model, MHC protects a peptide from further degradation; in addition, however, it also defines the identity of this peptide. Model 4 assumes that the polypeptide is degraded by specific endopeptidases independent of MHC. The resulting peptides are rapidly degraded by exopeptidases unless protected by binding to MHC. This model is essentially identical to a hypothesis formulated for the MHC class II restricted processing pathway¹⁴⁻¹⁶. Model 5 combines 2 and 4, model 6 combines 3 and 4. Models 5 and 6 would both explain the MHC-independent H-4^b prepeak as an intermediate peptide. If some peptides are not produced by proteolysis of complete proteins, but rather by local transcription or translation, as has been speculated¹⁸, the resulting peptides should be also influenced according to these models to explain MHC dependency of cellular peptide composition.

D^b-restricted H-Y peptide was present (Fig. 1j, k). The latter F₁ mouse expresses a mutant H-2K molecule (K^{bm1}) differing from K^b at three amino-acid residues and is otherwise identical to a (BALB/c × C57BL/6) F₁ mouse⁷. Thus, these data show that MHC class I molecules govern cellular peptide composition. The data also show that F₁ complementation occurs on the cellular level⁸ and on the peptide level, as extracts from both (BALB/c × C57BL/6)F₁ and (C57BL/6 × DBA/2)F₁ (Fig. 1l) (but neither parent, see Fig. 1b, c, m) contain the H-4^b main peak. Complementation is also seen for the D^b-restricted H-Y peptide, when the D^b gene is supplied only by the mother of a male mouse (Fig. 1l).

The small H-4^b prepeak around fraction 26 in Fig. 1a, b does not at first glance seem significant. But, this peak is highly reproducible (in 20 out of 23 preparations, Table 1) and reaches fairly high levels in some experiments (Fig. 1h, m). It is detected in extracts from H-4^b (or H-4^b-like) expressing cells also by the H-4^b-specific CTL clone B30X9-4a/34 (Fig. 2a, b, d), so that it should carry the same or similar antigenic determinants as the H-4^b main peak. The H-4^b prepeak can be detected in H-4^b (or H-4^b-like)-expressing mice of three different MHC haplotypes (H-2^b, H-2^d and H-2^a), including the strain B10.129-H-4^b, defining H-4, but not in H-2^k or H-2^b mice expressing H-4^a (Figs 2a-f, 1c). Thus, this prepeak represents an H-4^b-related peptide, whose presence in spleen cells is not dependent on polymorphic MHC molecules. Titration of BALB.B-derived fractions 26 (MHC-independent) and 29 (MHC-dependent) indicates 100-fold more efficient recognition of the main peak (Fig. 1d versus g).

Naturally processed minor histocompatibility peptides isolated from different mouse strains were rechromatographed by reversed-phase HPLC under conditions yielding high resolution. The H-4^b main peptides isolated from four mouse strains all elute at almost identical positions (44 ml; Fig. 2g). The same is

seen for H-Y peptides from three strains (42 ml) and for mapki peptides from two strains (36 ml; Fig. 3g). The small differences (± 0.5 ml) in different preparations are within the range of technical variability. Thus, it seems that cells of different genetic backgrounds but sharing a given MHC class I molecule produce and maintain identical peptides from a given protein. This high fidelity, however, might not be true for the MHC-independent H-4^b prepeak, which elutes with higher variations (33.5-35 ml).

We conclude that MHC class I gene products substantially influence the pool of peptides present in normal cells and that these peptides are produced by cells with high fidelity. The same appears to apply for cells transfected with a model CTL antigen, β -galactosidase (ref. 9 and data not shown), and for virus-infected cells¹⁰.

A commonly held view regarding MHC class I restricted antigen presentation is that nascent MHC class I molecules are loaded with pre-existing peptides in the endoplasmic reticulum^{5,11-13}. Our surprising result that the peptide pool is heavily dependent on MHC class I molecules does not exclude the possibility of pre-existing peptides independent of MHC, but indicates a substantial enrichment of MHC-binding peptides by MHC molecules, which is at least 500-fold for the H-4^b main peptide (in K^b-expressing cells versus K^b-negative cells; see Fig. 1). Thus, if a pool of pre-existing peptides independent of MHC exists in cells, it should be extremely short-lived, unless eventually protected by binding to MHC molecules¹⁴⁻¹⁶.

Protection of pre-existing peptides by MHC binding is not the only mechanism that could account for the observed MHC-dependency of cellular peptide patterns. We present six models resulting from an unbiased consideration on how MHC class I molecules could influence cellular peptide production and maintenance (Fig. 3) which should be useful in the design of experiments for dissection of the MHC class I-restricted antigen-processing pathway. □

Received 28 June; accepted 4 September 1990.

1. Bjorkman, P. J. et al. *Nature* **329**, 506-512 (1987).
2. Townsend, A. R. et al. *Cell* **44**, 959-968 (1986).
3. Wallny, H.-J. & Rammensee, H.-G. *Nature* **343**, 275-278 (1990).
4. Rötzschke, O., Falk, K., Wallny, H.-J., Faath, S. & Rammensee, H.-G. *Science* **249**, 283-287 (1990).
5. Townsend, A. et al. *Nature* **340**, 443-448 (1989).
6. Wettstein, P. J. in *Human Immunogenetics* (ed. Litwin, S. D.) 339-357 (Dekker, New York, 1989).
7. Klein, J., Figueroa, F. & David, C. S. *Immunogenetics* **17**, 553-596 (1983).
8. Bevan, M. J. *Nature* **256**, 417-421 (1975).
9. Rammensee, H.-G., Schild, H. & Theopold, U. *Immunogenetics* **30**, 296-302 (1989).
10. Rötzschke, O. et al. *Nature* **348**, 252-254 (1990).

11. Nuchtern, J. G., Bonifacio, J. S., Biddison, W. E. & Klausner, R. D. *Nature* **339**, 223-226 (1989).
12. Yewdell, J. W. & Bennink, J. R. *Science* **244**, 1072-1075 (1989).
13. Long, E. O. *Immun. Today* **10**, 232-234 (1989).
14. Werdell, O. *Scand. J. Immun.* **24**, 625-636 (1986).
15. Allen, P. M. & Unanue, E. R. A. *Inst. Pasteur Immun.* **137D**, 182 (1986).
16. Donermeyer, D. L., and Allen, P. M. *J. Immun.* **142**, 1063-1068 (1989).
17. Graff, R. J. in *Origins of Inbred Mice* (ed. Morse, H. C.) 371-389 (Academic, New York, 1978).
18. Boon, T. & Van Pel, A. *Immunogenetics* **29**, 75-79 (1989).

ACKNOWLEDGEMENTS. We thank J. Klein for support, P. Bohley for discussion, S. Faath for technical assistance, H.-J. Wallny for cell lines, H. Schild, H.-J. Wallny and P. Griem for reading and R. Kamarek for preparing the manuscript. Supported by Sonderforschungsbereich 120.

Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells

Olaf Röttzschke, Kirsten Falk, Karl Deres*,
Hansjörg Schild, Maria Norda, Jörg Metzger*,
Günther Jung* & Hans-Georg Rammensee†

Max-Planck-Institut für Biologie, Abteilung Immunogenetik,
Corrensstrasse 42, D-7400 Tübingen, Germany

* Institut für Organische Chemie, Universität Tübingen,
D-7400 Tübingen, Germany

† To whom correspondence should be addressed

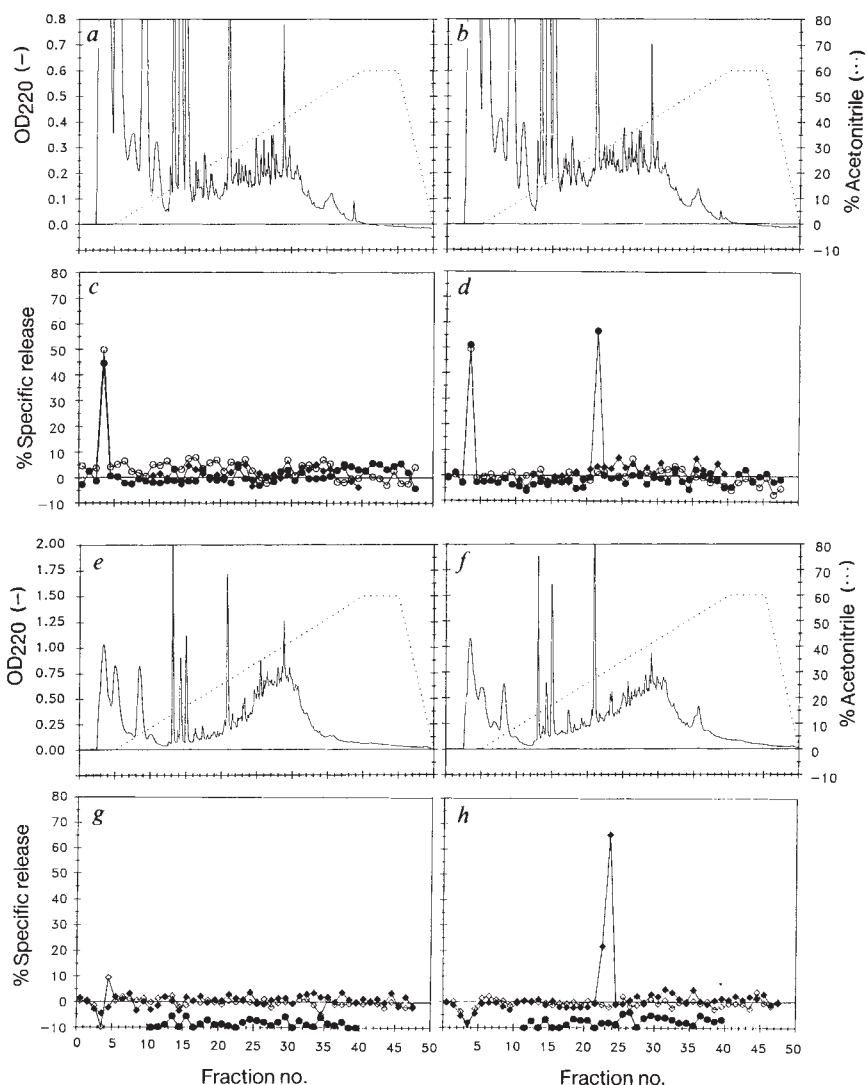
VIRUS-infected cells can be eliminated by cytotoxic T lymphocytes (CTL), which recognize virus-derived peptides bound to major histocompatibility complex (MHC) class I molecules on the cell surface^{1,2}. Until now, this notion has relied on overwhelming but indirect evidence, as the existence of naturally processed viral peptides has not been previously reported. Here we show that such peptides can be extracted from virus-infected cells by acid elution. Both the naturally processed H-2-D^b-restricted and H-2-K^d-

restricted peptides from influenza nucleoprotein are smaller than the corresponding synthetic peptides, which have first been used to determine the respective CTL epitopes^{1,3}. As with minor histocompatibility antigens⁴, occurrence of viral peptides seems to be heavily dependent on MHC class I molecules, because infected H-2^d cells do not contain the H-2-D^b-restricted peptide, and infected H-2^b cells do not contain the H-2-K^d-restricted peptide. Our data provide direct experimental proof for the above notion on MHC-associated viral peptides on virus-infected cells.

Influenza virus strain A/PR/8/34 was used to infect EL4 (H-2^b) or P815-TR (H-2^d) mouse tumour cells, respectively. The virus-infected cells were subjected to an acid elution procedure that has been used to isolate naturally processed minor histocompatibility antigens^{4,5}. Molecules with relative molecular masses less than 5,000 ($M_r < 5K$) contained in the resulting complex mixture were separated by reversed-phase HPLC (Fig. 1). Individual fractions were then tested for recognition by CTL specific for the dominant influenza virus epitopes, which are contained in amino-acid residues 365–380 from the nucleoprotein (NP; peptide NP365–380) for D^b-restricted CTL, and residues 147–158 from the same protein (NP147–158) for K^d-restricted CTL, respectively^{1,3}. Material recognized by D^b-restricted CTL eluted at fraction 22 of the extract from infected, but not uninfected, EL4 cells, whereas none of the fractions was recognized by K^d-restricted CTL (Fig. 1*a–d*). By contrast, D^b-

FIG. 1 Isolation of naturally processed viral CTL epitopes from virus-infected cells. EL4 (H-2^b) (*a–d*) or P815-TR (H-2^d) (ref. 12) (*e–h*) tumour cells (8×10^8) were left untreated (left-hand panels) or were infected with 50,000 units of influenza A/PR/8/34 virus (right-hand panels). Peptides were isolated from cells by acid extraction^{4,5} and separated by reversed-phase HPLC (*a, b, e, f*). Solid lines, absorption at 220 nm; dotted lines, percentage of acetonitrile in the gradient. *c, d, g, h*, Individual HPLC fractions were tested for recognition by influenza virus-specific CTL lines LS9 (D^b-restricted, specific for an epitope on NP365–380) (●) or HASI (K^d-restricted, specific for an epitope on NP147–158) (◆) or with medium (○, ◇). Target cells were EL4 cells (○, ●) for LS9 and P815 (◇, ◆) for HASI.

METHODS. Tumour cells were infected as described⁶. Cells were suspended in 0.1% trifluoroacetic acid (TFA), dounced, sonicated and centrifuged as described for spleen cells⁵. Material in the supernatant of $M_r > 5,000$ was removed by gel filtration (G25 Sepharose, Pharmacia). The remainder of the supernatant was separated on a reversed-phase HPLC column (Superpac Pep S, Pharmacia LKB) in 0.1% TFA using a gradient of increasing acetonitrile concentration. Flow rate, 1 ml min⁻¹; fraction size, 1 ml. Individual fractions were collected, dried, resuspended in PBS, incubated with ⁵¹Cr-labelled tumour cells (either EL4 or P815) and tested for recognition by influenza-specific CTL in a standard ⁵¹Cr release assay as described⁵. Spontaneous ⁵¹Cr release of target cells ranged between 15 and 25%. Effector to target ratio ranged between 5:1 to 20:1. CTL line HASI was produced by stimulating spleen cells of a BALB/c mouse (preimmunized with 50 units of A/PR/8/34 virus) with 100 ng ml⁻¹ of NP147–158 peptide in minimum essential alpha medium containing 10% FCS, β -mercaptoethanol, glutamine, and antibiotics at 37 °C, 5% CO₂, followed by weekly stimulation with irradiated (33 Gy) syngeneic spleen cells and peptide in medium supplemented with interleukin-2. This CTL line efficiently lyses A/PR/8/34-infected P815-TR, but not EL4 cells. The CTL line LS9 was produced by stimulating spleen cells of a (C57BL/6 $\times$ DBA/2)F1 mouse preimmunized with a synthetic lipopeptide vaccine⁸ containing NP365–380 according to the protocol used to produce HASI. LS9 efficiently lyses EL4 cells infected



with A/PR/8/34, but not infected P815-TR cells. Another D^b-restricted CTL line produced by immunization *in vivo* with virus showed recognition patterns of both natural and synthetic peptides identical to that of LS9 (not shown).

restricted CTL did not recognize any fraction from P815-TR extracts, whereas K^d-restricted CTL recognized fraction 24 from infected, but not from uninfected P815-TR cells (Fig. 1*e-h*). We conclude that naturally processed viral peptides exist, that they can be isolated from infected cells by acid elution, and that these peptides are involved in the MHC class I-restricted antigen-processing pathway, as suggested by their MHC dependency. The latter also excludes the possibility that the isolated peptides are artefacts produced during the extraction procedure. Because EL4 and P815-TR cells differ not only at MHC genes, the present data do not prove MHC dependency of processing ('processing' here means not only the cutting of proteins, but also the further fate of the degradation products). By use of MHC-recombinant and mutant mice, however, we have shown for minor histocompatibility antigens that the outcome of processing is dependent on MHC class I molecules⁴.

For their molecular identification, the biochemical behaviour of naturally processed peptides was compared with that of the corresponding synthetic peptides. Figure 2*a, b* shows reversed-phase HPLC profiles of preparations of synthetic peptides NP147-158 (TYQRTRALVRTG (single-letter amino-acid code)) and NP365-380 (IASNENMETMESSTLE), respectively. Recognition of individual fractions by the respective CTL is shown in Fig. 2*c, d*. Not only the main products were recognized, but also some byproducts of lower *M_r* (such byproducts in small amounts are common to crude synthetic peptide preparations). This result is consistent with reports that CTL may recognize shorter peptides as well, and sometimes much better than the above peptides (for example, NP366-379)^{1,6}. At higher dilution of fractions, D^b-restricted CTL failed to recognize the main peak, but still recognized some other fractions, which contained small to undetectable (according to absorbance at 220 nm) amounts of peptide (Fig. 2*f*). Incidentally, both crude synthetic peptide preparations NP147-158 and NP365-380 contained other peptides of smaller size, which coeluted exactly with the respective natural peptide (Fig. 2*e, f*). Thus, the naturally processed virus peptides in both cases are different,

most likely smaller, than the respective synthetic peptides reported^{1,3} to contain CTL epitopes. Both natural peptides are recognized in a concentration-dependent and MHC class I-restricted manner (Fig. 3*a, d*). The shorter synthetic by-products coeluting with the natural peptides, again in both cases, are recognized much better than are the nominal synthetic peptides (Fig. 3*b, e*). A truncated peptide in the crude synthetic NP147-158 preparation coeluting with the natural peptide was determined by ion spray mass spectrometry to be TYQRTRALV. A subsequently synthesized peptide according to this sequence (NP147-155) coeluted with the natural peptide (not shown) and was indeed recognized (used as crude material) much better than NP147-158 (Fig. 3*c*), about as well as is TYQRTRALVTG (missing R at position 156) which is recognized 1,000-fold better than NP147-158 (ref. 6; Fig. 3*c*). Thus the natural K^d-restricted CTL epitope of influenza NP is likely to be TYQRTRALV. Under this assumption, comparison of the titration curves in Fig. 3*a, b, c* allows the calculation of the upper limit of peptide molecules extracted per infected cell to be 1,000. By similar analysis, the natural D^b-restricted peptide coeluted with ASNENMETM (NP366-374), which is recognized 1,000 times better than IASNENMETMESSTLE (manuscript in preparation). Thus, a previously identified optimal D^b-restricted NP peptide (NP366-379) happens to share its N-terminal amino acid residue with the natural peptide¹.

Together our experiments show that virus-infected cells produce small peptides from viral proteins which are recognized by MHC class I-restricted CTL, thereby directly confirming the overwhelming evidence provided by MHC crystallography and by experiments with synthetic peptides and truncated genes^{1,2,7}. The data also indicate that the use of synthetic peptides to identify T-cell epitopes may be misleading, as very minor byproducts may be responsible for much of the biological effect.

The results in this and in a previous paper on naturally processed minor histocompatibility peptides⁴ show two features of MHC class I-restricted antigen presentation. First, a cell produces and maintains exactly one peptide presented to a given

FIG. 2 Naturally processed influenza virus-derived CTL epitopes coelute with synthetic peptides. *a*, Reversed-phase HPLC (at higher resolution as in Fig. 1) elution profiles (solid lines) of synthetic peptide preparations according to NP147-158 (*a*) or NP365-380 (*b*). Dashed lines, percentage of acetonitrile in the gradient. *c*, Recognition of HPLC fractions (diluted 1:2,500) of the NP147-158 preparation by K^d-restricted CTL HAS1 (◆) or without CTL (◇) using P815 as target cells. *d*, Recognition of fractions of the synthetic NP365-380 preparation at same dilution by D^b-restricted CTL LS9 (●) or without CTL (○) using EL4 as target cells. *e*, Recognition of NP147-158 fractions (diluted 1:62,500) (◆) and of rechromatographed natural K^d-restricted peptide (fraction 24 of Fig. 1*f, h*) (▼) by HAS1 CTL. *f*, Recognition of NP365-380 fractions (diluted 1:62,500) (●) and of rechromatographed natural D^b-restricted peptide (fraction 22 of Fig. 1*b, d*) (▲) by LS9 CTL.

METHODS. Peptides were synthesized as described⁸. Peptide preparations (400 µg of each) were chromatographed on the same reversed-phase HPLC column as in Fig. 1, using conditions yielding higher resolution (slower increase of acetonitrile concentration). Flow rate, 1 ml min⁻¹; fraction size, 0.5 ml. The main peaks in *a* and *b* were confirmed by ion-spray tandem-mass-spectrometry to be identical with the nominal peptide sequences, respectively. The fractions identified as naturally processed viral antigens in Fig. 1 were rechromatographed using exactly the same conditions as for the synthetic peptide preparations. Individual fractions were tested for CTL recognition as in Fig. 1.

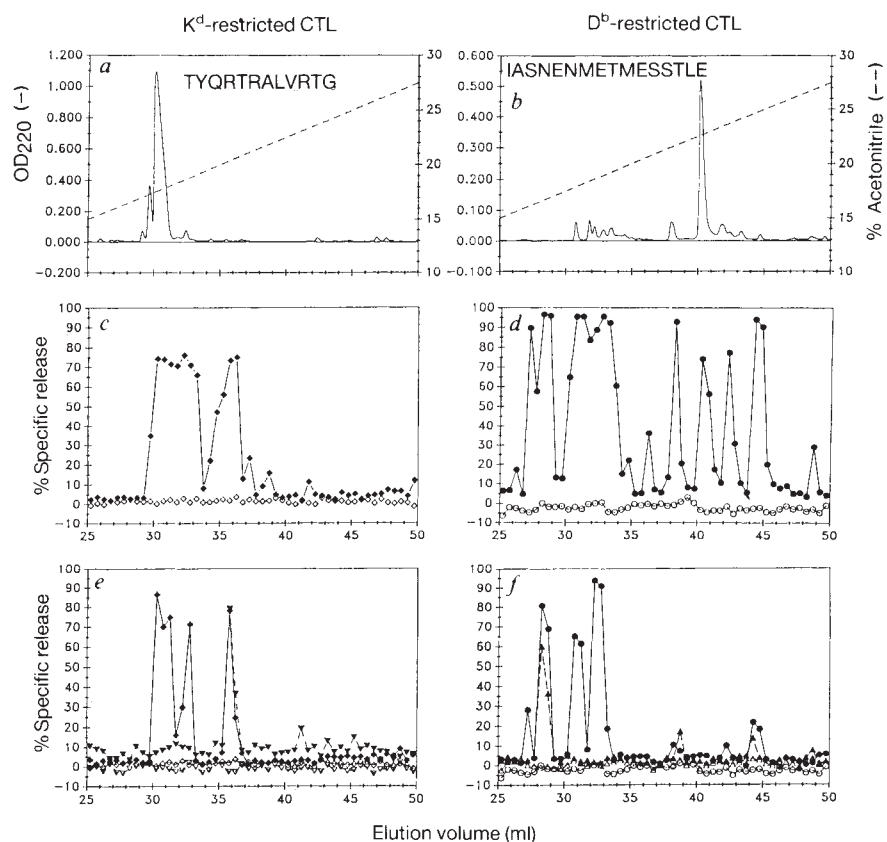
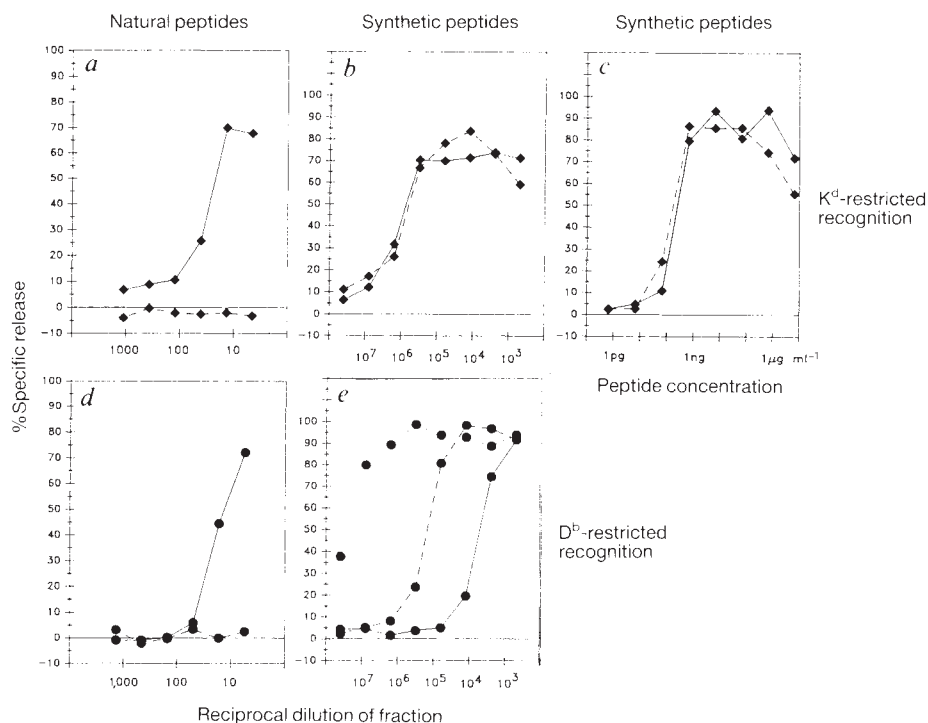


FIG. 3 Titration of natural and synthetic peptide fractions. *a*, Natural K^d -restricted NP peptide (fraction 24 of Fig. 1*f, h*) was tested in titrated concentrations for recognition by HASI CTL using P815 (—) or EL4 (---) as target cells. No killing of EL4 indicates MHC class I-restricted recognition. *b*, Fractions of Fig. 2*a* representing the main synthetic peptide peak (TYQRTRALVRTG; 30.5–31.0 ml elution volume) (—) and the truncated peptide (TYQRTRALV according to tandem-mass-spectrometric analysis) coeluting with the natural peptide (35.5–36.0 ml) (---) were tested in titrated dilution steps for recognition by HASI CTL. Note the difference between the peptide concentrations of the two samples, as evident from their absorption values in Fig. 2*a*. *c*, Crude batches of synthetic peptide preparations TYQRTRALV (---) (synthesized on a fully automated simultaneous multiple peptide synthesizer; Zinsser Analytical model 350)¹² and TYQRTRALVTG (—) (missing R at position 156)^{6,8} were tested in 10-fold dilution steps for recognition by HASI CTL. *d*, Natural D^b -restricted NP peptide (fraction 22 of Fig. 1*b, d*) was titrated into a CTL assay with LS9 CTL using EL4 (—) or P815 (---) as targets. *e*, Fractions of Fig. 2*b* representing the main synthetic peptide (IASNENMETMESSTLE; 40.0 to 40.5 ml elution volume) (—), the truncated synthetic peptide coeluting with the natural peptide (28.0–28.5 ml)



CTL, as shown by elution of the natural peptide as a single sharp activity peak on reversed-phase HPLC profiles. Second, the processing of endogenous proteins seems to be dependent on MHC class I molecules, as discussed in ref. 4. A database containing many sequences of naturally processed peptides should allow insight into the specificity of the proteases involved, and further our knowledge about the rules of peptide-MHC interactions.

The acid elution method for isolation of naturally processed T-cell epitopes in combination with peptide chemistry including highly sensitive analytical tools such as HPLC and capillary-zone electrophoresis combined with ion-spray tandem-mass-spectrometry should provide straightforward experimental access to the peptide sequences of processed pathogen-derived antigens and has impact on synthetic vaccine design: the efficiency of synthetic peptide vaccines for CTL activation^{8,9} could be improved further by using exactly the peptide produced by the infected cells. On the other hand, knowing exactly which peptide is presented naturally to a given T cell could help in the design of prophylactic or therapeutic measures for T cell-mediated autoimmune diseases¹⁰. □

(---), and a second truncated, hyper-reactive peptide (32.0–32.5 ml) (· · ·) were tested in 10-fold dilution steps for recognition by LS9 CTL.

Three-dimensional structure of an idiotope–anti-idiotope complex

G. A. Bentley, G. Boulout, M. M. Riottot & R. J. Poljak

Unité d'Immunologie Structurale, URA 359 CNRS,
Département d'Immunologie, Institut Pasteur,
75724 Paris Cedex 15, France

SEROLOGICALLY detected antigenic determinants unique to an antibody or group of antibodies^{1–3} are called idiotopes. The sum of idiotopes of an antibody constitute its idiotype³. Idiotypes have been intensively studied following a hypothesis for the self-regulation of the immune system through a network of idiotype–anti-idiotype interactions⁴. Furthermore, as antigen and anti-idiotypes can competitively bind to idiotope-positive, antigen-specific antibodies, anti-idiotypes may carry an 'internal image' of the external antigen (see refs 5–10 for reviews). Here we describe the structure of the complex between the monoclonal anti-lysozyme FabD1.3 and the anti-idiopathic FabE225 at 2.5 Å resolution. This complex defines a private idiotope consisting of 13 amino-acid residues, mainly from the complementarity-determining regions of D1.3. Seven of these residues make contacts with the antigen, indicating a significant overlap between idiotope and antigen-combining site. Idiopathic mimicry of the external antigen is not achieved at the molecular level in this example.

The crystal structure of the idiotope–anti-idiotope complex FabD1.3–FabE225 (ref. 11) shows the two Fabs roughly aligned along their major lengths, interacting largely through their complementarity-determining regions (CDRs) (Fig. 1*a*). Fourteen residues from the six CDRs of E225, together with a framework residue of V_L , contribute to interatomic contacts with the idiotope (Table 1). In turn, five CDRs and one V_L framework loop contribute 13 residues to the idiotope of D1.3 recognized by E225 (Table 2). FabE225 is centred on V_L of FabD1.3 such that nine residues of the idiotope are located on this domain

Received 10 July; accepted 3 October 1990.

1. Townsend, A. R. *et al.* *Cell* **44**, 959–968 (1986).
2. Bjorkman, P. J. *et al.* *Nature* **329**, 506–512 (1987).
3. Taylor, P. M., Davey, J., Howland, K., Rothbard, J. B. & Askonas, B. A. *Immunogenetics* **26**, 267–272 (1987).
4. Falk, K., Rötzschke, O. & Rammensee, H.-G. *Nature* **348**, 248–251 (1990).
5. Rötzschke, O., Falk, K., Wallny, H.-J., Faath, S. & Rammensee, H.-G. *Science* **249**, 283–287 (1990).
6. Bodmer, H. C., Pemberton, R. M., Rothbard, J. B. & Askonas, B. A. *Cell* **52**, 253–258 (1988).
7. Gould, K., Cossins, J., Bastin, J., Brownlee, G. G. & Townsend, A. *J. exp. Med.* **170**, 1051–1056 (1989).
8. Deres, K., Schild, H., Wiesmüller, K.-H., Jung, G. & Rammensee, H.-G. *Nature* **342**, 516–564 (1989).
9. Aichele, P., Hengartner, H., Zinkernagel, R. M. & Schulz, M. *J. exp. Med.* **171**, 1815–1820 (1990).
10. Wraith, D. C., Smilek, D. E., Mitchell, D. J., Steinman, L. & McDevitt, H. O. *Cell* **59**, 247–255 (1989).
11. Van Pel, A., De Plaen, E. & Boon, T. *Somatic Cell molec. Genet.* **11**, 467–475 (1985).
12. Schnorrenberg, G. & Gerhardt, H. *Tetrahedron* **45**, 7759–7764 (1989).

ACKNOWLEDGEMENTS. We thank J. Klein for support, S. Faath for technical assistance, A. Vallbracht for help with the virus preparation, H.-J. Wallny, P. Griem and K.-H. Wiesmüller for critically reading, and R. Kamarek for preparing the manuscript. K.D. is supported by Studienstiftung des Deutschen Volkes. This research is supported by Sonderforschungsbereich 120 and 323.

TABLE 1 Contacts of E225 (anti-idiotope) with D1.3 (idiotope)

Light chain contacts					Heavy chain				
Contacts	E225	D1.3	No. contacts	H-bond or salt bridge	Contacts	E225	D1.3	No. contacts	H-bond or salt bridge
L1	R30	W352	7	1 salt bridge	H1	W33	N31	5	1 H-bond
		D354	4		H2	Y52	T52	3	
		Y401	1			S55	S65	1	
		Y401	1			D57	G66	5	
L2	W50	Y32	3	1 H-bond	H3	N59	N31	1	1 H-bond
		Y50	7			L100	Y50	3	
		D400	6			F102	H30	2	
		Y401	5				N31	4	
FR3	S67	D354	6	1 H-bond			Y32	10	
L3	C91	Y32	2	1 H-bond			Y50	4	
		Y32	1						
		W92	2						
		H30	2						
	S93	W92	1						
Y94	H30	H30	10	1 H-bond					
		N31	3						

L1, L2, L3, H1, H2 and H3 are the CDRs 1, 2 and 3 of the L and H chains, respectively. FR3 refers to the third framework region²⁰. The one-letter code is used for amino acids; their sequential numbers are from 1 to 214 for the L-chain and from 301 to 518 for the H-chain of D1.3. Contacts are hydrogen bonds, salt bridges or van der Waal's interactions for which the maximum interatomic distances in Å are: C-C, 4.11; N-N, 3.44; O-O, 3.33; C-N, 3.77; C-O, 3.72; N-O, 3.39 (refs 15, 21); the maximum interatomic distance for the electronegative atoms of a hydrogen-bond pair is 3.5 Å. The boxed residues of the idiotope (D1.3) are those which contact¹⁶ the external antigen HEL.

Similarly, FabD1.3 is centred on V_L of the anti-idiotope (Fig. 1b); however, 45% of the contacts are made by V_H E225. About 800 Å² of solvent-accessible surface in each Fab is buried on complex formation.

Polar interactions in the complex include one salt bridge formed between the anti-idiotope V_L Arg 30 and V_H Asp 54 of the idiotope, as well as nine hydrogen bonds (Table 1). The V_L framework residue Ser 67 of E225, situated at the extremity of a loop adjacent to the CDR2, also contacts the idiotope. Although this loop is outside the conventional combining site for a globular protein antigen, the more unusual topology offered by a multidomain antigen as FabD1.3 allows it to participate in antigen binding. Similarly, the V_L residues Ser 65, Gly 66 and Ser 67, from the homologous loop of D1.3, make contacts with E225, including a hydrogen bond by Ser 65. Thus, it is conceivable that the homologous framework loop of the T-cell receptor could interact with antigen.

The individual idiotope of D1.3 recognized by monoclonal antibody E225 consists of amino-acid residues belonging to the V_H and V_L CDRs, as well as framework residues of a nearby loop of V_L (Ser 65, Gly 66 and Ser 67). Thirteen amino-acid residues contribute to the idiotope, a number which is similar to that of antigenic determinants that have so far been described¹²⁻¹⁵. However, the idiotope comprises amino-acid residues from six discontinuous stretches of the polypeptide chains of the V_H and V_L domains (Table 1). Thus, it is a topographical, folding-dependent antigenic determinant. In addition, the highly variable CDRs of V_H and V_L make this determinant a structurally unique, or 'private', idiotope.

The lysozyme (HEL)-binding site (or paratope) of D1.3, now well-defined from the refined crystal structures of the free Fv and the Fv-HEL and Fab-HEL complexes¹⁶, consists of a set of 13 residues (or 14 if contacts through bridging water molecules are considered). Of the 13 residues comprising the idiotope of D1.3 recognized by E225, 7 are in common with the paratope (Tables 1, 2). Structural comparison of the CDRs of D1.3 in the idiotope-anti-idiotope complex with those of the Fv and Fab antigen complexes shows that no important difference exists either between their main-chain conformations or between their relative disposition in space. We find, however, that the side-

chain conformation of three CDR residues, shared between the paratope and idiotope, differ significantly. The observed differences can be explained by the different steric requirements for the binding of HEL and E225 (Fig. 1c; Table 3). We note, by contrast, that no significant difference in the CDR residues, either in the main-chain or side-chain conformations, is seen between the HEL-bound and free states of the Fv fragment of D1.3 (ref. 16).

Because the HEL epitope recognized by monoclonal antibody D1.3 is partly in α -helical conformation, it may be difficult to mimic by the open loop structures of the CDRs of antibodies. Furthermore, as overlap between the paratope and idiotope of D1.3 is only partial, the potential for total molecular mimicry of HEL by E225 is reduced. Indeed, a detailed analysis of the overlap (Fig. 1b, c) indicates that E225 does not provide, at an

TABLE 2 Contacts of D1.3 (idiotope) with E225 (anti-idiotope)

D1.3		E225			
L1	H30	S93(2)	Y94(10)	F402(2)	
	N31	Y94(3)	W333(5)	N359(1)	F402(4)
	Y32	W50(3)	C91(2)	G92(1)	F402(10)
L2	Y50	W50(7)	L400(3)	F402(4)	
	T52	W333(3)	Y352(8)		
FR3	S65	S355(1)			
	G66	D357(5)			
	S67	D357(3)			
L3	W92	G92(2)	S93(1)		
H2	W52	R30(7)			
	D54	R30(4)	S67(6)		
H3	D100	W50(6)			
	Y101	R30(1)	I31(1)	W50(5)	

L1, L2, L3, FR3, H2 and H3 as in Table 1. One-letter code used for amino acids. The sequential numbers for E225 are from 1 to 214 and from 301 to 519 for residues of the L- and H-chains, respectively. Boxed residues of the idiotope are those that contact the external antigen HEL. The value in parenthesis after each E225 residue corresponds to the number of contacts (as defined in Table 1) made by that residue.

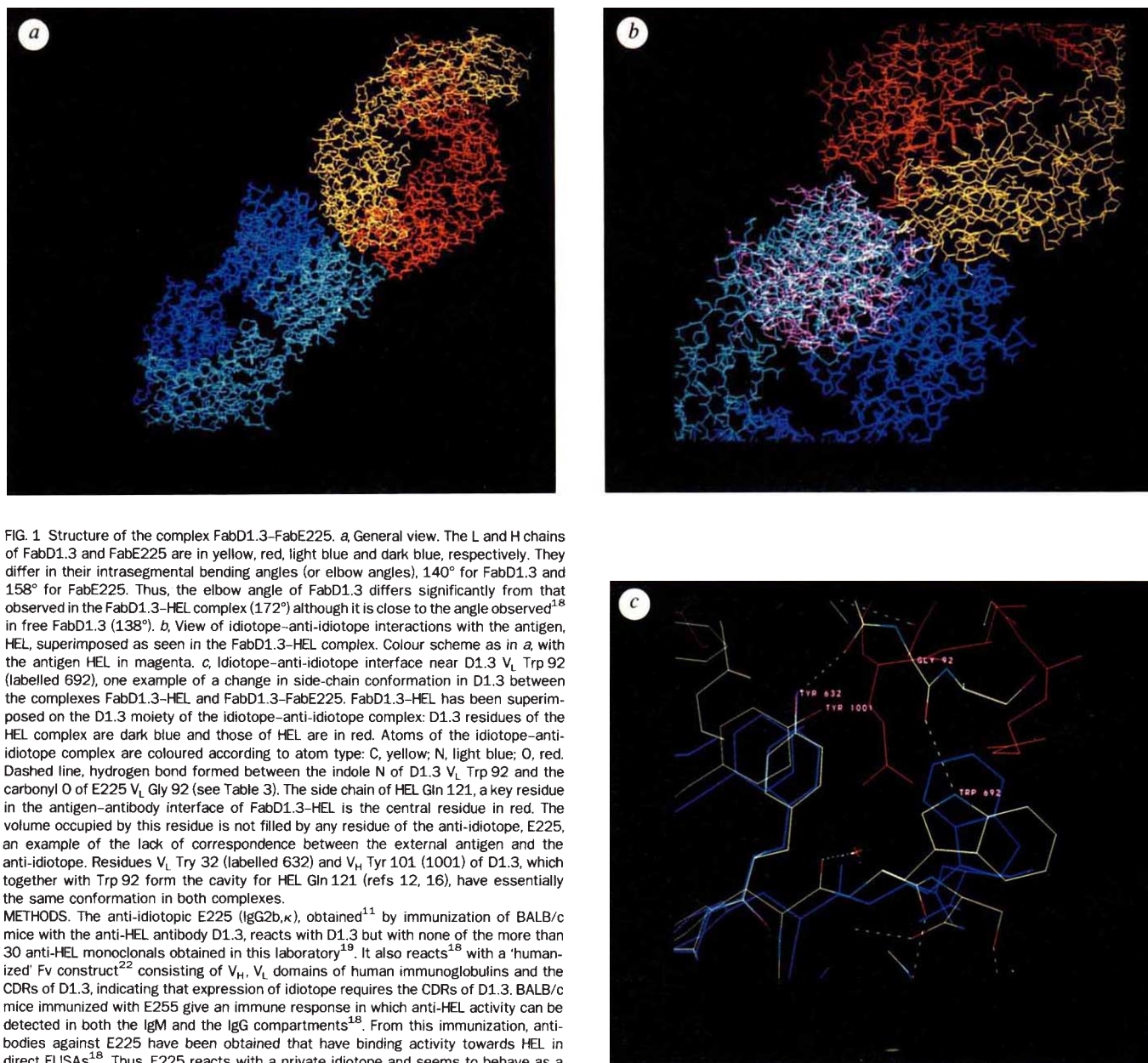


FIG. 1 Structure of the complex FabD1.3-FabE225. *a*, General view. The L and H chains of FabD1.3 and FabE225 are in yellow, red, light blue and dark blue, respectively. They differ in their intrasegmental bending angles (or elbow angles), 140° for FabD1.3 and 158° for FabE225. Thus, the elbow angle of FabD1.3 differs significantly from that observed in the FabD1.3-HEL complex (172°) although it is close to the angle observed¹⁸ in free FabD1.3 (138°). *b*, View of idiotope-anti-idiotope interactions with the antigen, HEL, superimposed as seen in the FabD1.3-HEL complex. Colour scheme as in *a*, with the antigen HEL in magenta. *c*, Idiotope-anti-idiotope interface near D1.3 V_L Trp 92 (labelled 692), one example of a change in side-chain conformation in D1.3 between the complexes FabD1.3-HEL and FabD1.3-FabE225. FabD1.3-HEL has been superimposed on the D1.3 moiety of the idiotope-anti-idiotope complex: D1.3 residues of the HEL complex are dark blue and those of HEL are in red. Atoms of the idiotope-anti-idiotope complex are coloured according to atom type: C, yellow; N, light blue; O, red. Dashed line, hydrogen bond formed between the indole N of D1.3 V_L Trp 92 and the carbonyl O of E225 V_L Gly 92 (see Table 3). The side chain of HEL Gln 121, a key residue in the antigen-antibody interface of FabD1.3-HEL is the central residue in red. The volume occupied by this residue is not filled by any residue of the anti-idiotope, E225, an example of the lack of correspondence between the external antigen and the anti-idiotope. Residues V_L Trp 32 (labelled 632) and V_L Tyr 101 (1001) of D1.3, which together with Trp 92 form the cavity for HEL Gln 121 (refs 12, 16), have essentially the same conformation in both complexes.

METHODS. The anti-idiotopic E225 (IgG2b, κ), obtained¹¹ by immunization of BALB/c mice with the anti-HEL antibody D1.3, reacts with D1.3 but with none of the more than 30 anti-HEL monoclonals obtained in this laboratory¹⁹. It also reacts¹⁸ with a 'humanized' Fv construct²² consisting of V_H , V_L domains of human immunoglobulins and the CDRs of D1.3, indicating that expression of idiotope requires the CDRs of D1.3. BALB/c mice immunized with E255 give an immune response in which anti-HEL activity can be detected in both the IgM and the IgG compartments¹⁸. From this immunization, antibodies against E225 have been obtained that have binding activity towards HEL in direct ELISAs¹⁸. Thus, E225 reacts with a private idiotope and seems to behave as a member of the 'internal image' set^{4,17} of anti-D1.3 anti-idiotopes. Structure determination: The crystals are monoclinic $P2_1$, with $a = 75.7$ Å, $b = 77.4$ Å, $c = 97.2$ Å, $\beta = 112^\circ$, and one complex per asymmetric unit. X-ray intensities were measured at the Multiwire Area Detector Facility²³, University of Virginia: 33,526 independent native reflexions to 2.5 Å resolution (97% complete); 20,668 independent reflexions to 2.9-Å resolution for a $K_3UO_2F_5$ derivative; 6,826 independent reflexions to 4.3-Å resolution for a tetrakis-(acetoxymethyl)-methane derivative. The structure was determined by combination of multiple isomorphous replacement and molecular replacement methods. Both heavy-atom derivatives suffered from non-isomorphism and phases were not determined beyond 4-Å resolution (average figure merit, 0.53). In parallel, molecular replacement was applied using the variable (V_H and V_L) and constant (C_H1 and C_L) domains

of FabD1.3 as search models. Rotation and translation functions gave orientations and positions of two variable and two constant dimers as required for two Fabs in the asymmetric unit. These were referred to a common unit-cell origin using the isomorphous replacement phase information in the phased translation function²⁴. Refinement was performed using the programs CORELS²⁵, PROLSQ²⁶ and X-PLOR²⁷, the sequence of E225 has been published²⁸. About 80 solvent molecules were placed. The final R factor was 17.9% with a r.m.s. difference from ideal bond lengths and bond angles of 0.016 Å and 2.4° . Assessment²⁹ of the intrinsic accuracy of atomic coordinates gives a mean error of 0.3 Å. Further details of analysis will be published elsewhere.

atomic level, an internal image of the external antigen HEL. The simple qualitative comparison provided in Table 3 shows that the nature of the interaction of a D1.3 residue is often different between the two complexes FabD1.3-HEL and FabD1.3-FabE225. This is due in part to structural differences occurring at the combining site of D1.3 between the antigen bound¹⁶ and the anti-idiotope-bound state described above. More generally, therefore, accurate antigen mimicry may not be achieved at the atomic level as the antigen-binding template

is adaptable to the different ligands, antigen or anti-idiotope with which it is confronted. In this context it is worth considering that although anti-idiotopic antibodies seem in many cases to activate biological receptors (see ref. 17 for review), this may be due to micro-aggregation of the receptors. Since bivalent antibodies (but not their Fab fragments) will facilitate aggregation of target molecules, they may be able to activate receptors even with a less specific binding than that effected by physiological ligands. Receptors that share the immunoglobulin fold or

TABLE 3 D1.3 residues in common with paratope and idiotope

	D1.3	D1.3-HEL	D1.3-E225
V_L	Y32	v.d.w.	1 H-bond
	Y50	1 H-bond	1 H-bond
	W92	v.d.w.	1 H-bond
V_H	W52	v.d.w.	v.d.w.
	D54	1 H-bond	1 salt bridge, 1 H-bond
	D100	5 H-bonds	v.d.w.
	Y101	1 H-bond	v.d.w.

Residues shared by the idiotope and the paratope are listed in the column under D1.3. Their contacts with the antigen¹⁶ and with the anti-idiotope are listed under D1.3-HEL and D1.3-E225, respectively. The nature of the interactions are designated as v.d.w. for van der Waals's contacts and H-bond for hydrogen bonds. Three of these residues have significant differences in side chain conformations. (1) V_L Tyr 50 has turned by $\sim 150^\circ$ about the C α -C β bond to avoid steric hindrance that would otherwise occur in contacting E225. (2) The ring of V_L Trp 92 is rotated by close to 180° such that its indole N makes a hydrogen bond with the carbonyl O of V_L E225 Gly 92 (see Fig. 1c). By contrast, this residue makes hydrophobic contacts¹⁶ in its complex with the external antigen, HEL. (3) The carboxyl group of Asp 100 differs by a rotation of about 90° as steric hindrance would not allow the conformation adopted in the complex with HEL. These structural changes are an example of an induced fit in the contacts made by the idiotope of antibody D1.3. Apart from these differences, the side-chain conformations of the CDR residues remain strikingly close to those observed in the FabD1.3-HEL, FvD1.3-HEL and FvD1.3 structures¹⁶.

conformationally flexible peptides may provide more favourable cases for anti-idiotopic mimicry.

In conclusion, although the immune system can be harnessed to provide complementary structures, the use of anti-idiotopic antibodies may not always provide exact images of external antigens at the molecular level. It is likely that the practical use of this approach will decrease with the increasing complexity of the target antigens, from oligopeptides to fully folded, multi-subunit proteins. The anti-idiotopic approach will be less adequate when high specificity and affinity are required, as, for example, in mimicking vaccinating antigens to confer immune protection. $\square$

Received 12 July; accepted 5 September 1990.

- Lohss, F., Weiler, E. & Hillmann, G. Z. *Naturforschung* **58**, 625-631 (1953).
- Slater, R. J., Ward, S. M. & Kunkel, H. G. J. *exp. Med.* **101**, 85-108 (1955).
- Oudin, J. & Michel, M. C. R. *hebd. Acad. Sci., Paris* **257**, 805-808 (1963).
- Jerne, N. K. A. *Immun. Inst. Pasteur* **125C**, 373-389 (1974).
- Urbain, J., Wulimart, C. & Cazenave, P.-A. *Contemp. Top. molec. Immun.* **8**, 113-148 (1981).
- Kabat, E. A. in *The Biology of Idiotypes* (eds Greene, M. I. & Nisonoff, A.) 3-17 (Plenum, New York, 1984).
- Erlanger, B. F. *Immun. Today* **10**, 151-152 (1989).
- Nisonoff, A. & Lamoyi, E. *Clin. Immun. Immunopath.* **21**, 397-406 (1981).
- Roitt, J. M. et al. *Lancet* **1**, 1041-1045 (1981).
- Sacks, D. L., Kelsoe, G. H. & Sachs, D. H. *Springer Semin. Immunopath* **6**, 79-97 (1983).
- Boulton, G. et al. *J. molec. Biol.* **194**, 577-579 (1987).
- Amit, G. A., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. *Science* **233**, 747-753 (1986).
- Sheriff, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8075-8079 (1987).
- Colman, P. et al. *Nature* **326**, 358-363 (1987).
- Tulip, W. et al. *Cold Spring Harbor Symp. quant. Biol.* **54**, 257-263 (1990).
- Bhat, T. N., Bentley, G. A., Fischmann, T., Boulton, G. & Poljak, R. J. *Nature* **347**, 483-485 (1990).
- Gaulton, G. N. & Greene, M. I. A. *Rev. Immun.* **4**, 253-280 (1986).
- Bentley, G. A. et al. *Cold Spring Harbor Symp. quant. Biol.* **54**, 239-245 (1990).
- Harper, M., Lema, F., Boulton, G. & Poljak, R. J. *Molec. Immun.* **24**, 97-108 (1987).
- Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. & Gottesman, K. S. *Sequences of Proteins of Immunological Interest* 4th ed (US Dept. Health and Public Services, 1987).
- Sheriff, S., Hendrickson, W. A. & Smith, J. L. J. *J. molec. Biol.* **197**, 273-296 (1987).
- Verhoeyen, M., Milstein, C. & Winter, G. *Science* **239**, 1534-1536 (1988).
- Sobotka, S. F. et al. *Nucl. Inst. Meth.* **220**, 575-581 (1984).
- Colman, P., Fehlihammer, H. & Bartels, K. in *Cryst. Computing Tech.* (eds Ahmed, F. R., Hulm, K. & Sedlaczek, B.) 248-258 (1976).
- Sussman, J. *Meth. Enzym.* **115**, 271-303 (1985).
- Hendrickson, W. *Meth. Enzym.* **115**, 253-270 (1985).
- Brünger, A., Kurian, J. & Karplus, M. *Science* **235**, 458-460 (1987).
- Souchon, H., Doyen, N., Riottot, M. M., Rougeon, F. & Poljak, R. J. *Molec. Immun.* **27**, 429-433 (1990).
- Luzzati, V. *Acta Cryst.* **5**, 802-810 (1952).

ACKNOWLEDGEMENTS. Supported by grants from the Institut Pasteur, CNRS, Fondation L. Jeantet, US Public Health Service and the EEC. We thank R. J. Chandross, R. H. Kretsinger and S. Sobotka for advice during X-ray intensity data collection, T. Fischmann and T. N. Bhat for discussions.

Structure, expression and function of a schwannoma-derived growth factor

Hideo Kimura, Wolfgang H. Fischer & David Schubert

The Salk Institute for Biological Studies,
PO Box 85800, San Diego, California 92186-5800, USA

DURING the development of the nervous system, cells require growth factors that regulate their division and survival. To identify new growth factors, serum-free growth-conditioned media from many clonal cell lines¹ were screened for the presence of mitogens for central nervous system glial cells. A cell line secreting a potent glial mitogen was established from a tumour (or 'schwannoma') derived from the sheath of the sciatic nerve. The cells of the tumour, named JS1 cells, were adapted to clonal culture and identified as Schwann cells. Schwann cells secrete an autocrine mitogen² and human schwannoma extracts have mitogenic activity on glial cells³. Until now, neither mitogen has been purified. Here we report the purification and characterization of a mitogenic molecule, designated schwannoma-derived growth factor (SDGF), from the growth-conditioned medium of the JS1 Schwann cell line. SDGF belongs to the epidermal growth factor family, and is an autocrine growth factor as well as a mitogen for astrocytes, Schwann cells and fibroblasts.

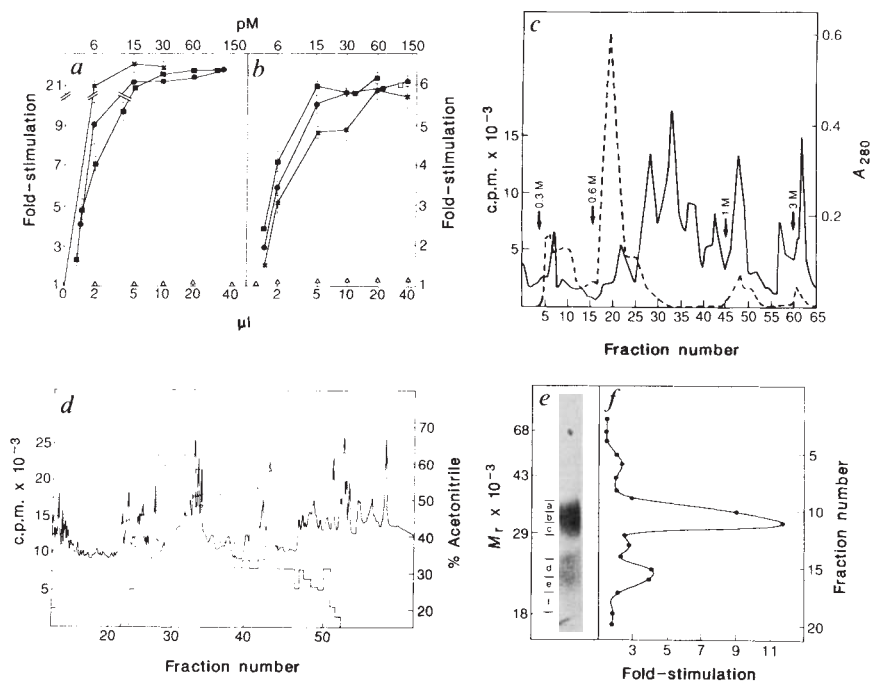
When serum-free growth-conditioned medium of JS1 cells is applied to Swiss 3T3 cells and primary central nervous system (CNS) astrocytes, it stimulates the incorporation of [³H]TdR into DNA. (Fig. 1a, b and Table 1). The JS1 mitogen binds to heparin-Sepharose and elutes from the column at 0.6 M sodium chloride (Fig. 1c). After reverse phase chromatography (Fig. 1d), the active HPLC fractions contain a main broad band on SDS-polyacrylamide gels with an apparent relative molecular mass of between 31,000 and 35,000 (M_r 31-35K) (Fig. 1e) which has most of the mitogenic activity (Fig. 1f). A smaller amount of mitogenic protein of M_r 20-25K is also present. To determine a partial amino-acid sequence of the mitogens, these proteins were electrophoretically transferred to Immobilon and excised for N-terminal sequence determination. Identical amino-acid sequences were obtained from several fractions of the main broad band as well as the lower M_r form of the protein (Fig. 1, legend).

Two 32-fold degenerate 17-mer deoxyoligonucleotides, designated JN1 and JN2, complementary to sequences encoding the N-terminal amino acids 1-6 and 8-13, were synthesized, and a λ gt10 library from JS1 complementary DNA was screened using the JN1 probe. Screening 10⁵ clones yielded 10 positive clones, one of which also hybridized with the JN2 probe. Nucleotide sequencing of the 1.2-kilobase (kb) insert revealed a 243-residue prepro-protein and a 3'-untranslated sequence (Fig. 2a). The ATG codon at position 1-3 precedes a basic amino acid (Arg) followed by a signal peptide of 19 hydrophobic residues. A stop codon (TAG) is present at nucleotides 730-732, and a polyadenylation signal (ATTAAA) 17 nucleotides upstream from the poly(A) sequence is found at the 3' end. The N-terminal sequence of the purified SDGF is at amino acid 97, indicating that a propeptide is cleaved from the precursor to generate a mature protein. It is not known whether SDGF undergoes proteolytic processing in the C terminus, but as lower M_r forms of SDGF have the same N-terminal sequence (Fig. 1), this is a distinct possibility. Although the preproprotein has a predicted M_r 26,633 and the mature protein a predicted M_r of 16,709, the observed M_r of the most abundant form of the protein is between 30 and 35K (Fig. 1e). This discrepancy is probably due to an N-linked glycosylation site in addition to several potential O-linked glycosylation sites.

SDGF is 76% homologous to human amphiregulin⁴⁻⁶, a

FIG. 1 Mitogenic activities and purification of SDGF. *a*, Swiss 3T3 fibroblasts (obtained from T. Hunter (Salk Institute)), and the effect of the various JS1 fractions on [3 H]-TdR incorporation assayed as described previously⁹. Data are presented as the mean of the fold-stimulation of [3 H]-TdR incorporation (no stimulation is given as 1 and a doubling of incorporation is 2) in quadruplicate wells $\pm$ s.e.m. versus protein or medium concentration. Basal level [3 H]-TdR incorporation was 650 c.p.m. *b*, Primary cultures of cortical astrocytes were prepared according to McCarthy and deVellis¹⁰ and stimulation of [3 H]-TdR incorporation determined as in *a*. Astrocyte cultures were >95% positive for glial fibrillary acidic protein. Basal incorporation 1,500 c.p.m. For transfections with SDGF cDNA, 293 human kidney cells were grown on 100-mm plates in DME supplemented by 10% FCS. Medium was changed to fresh DME with 10% FCS 3 h before transfection. DNA of SDGF expression plasmid¹¹ was mixed with calcium chloride and HEPES (pH 7.3) buffer and transfection carried out as described¹¹. The cells were then incubated in serum-free DME for 48 h under 3% CO₂. The resultant supernatants were assayed for mitogenic activity. $\times$, SDGF-transfected 293, supernatant; Δ , control-transfected 293, supernatant; $\square$, total JS1 serum-free growth-conditioned medium; $\bullet$, HPLC-purified SDGF. Protein concentration was determined by amino-acid analysis.

METHODS. Serum-free growth-conditioned medium was prepared by growing JS1 cells to near confluence in DME containing 10% horse serum in 100 mm tissue culture dishes. The cells were washed twice with serum-free DME and covered with 3 ml of serum-free DME. Two to three days later the medium was collected, centrifuged at 15,000 r.p.m. for 10 min to remove debris, and frozen at -20°C . To purify JS1 mitogen, 4 litres of serum-free JS1 growth-conditioned medium was directly passed through a 2 cm $\times$ 15 cm heparin-Sepharose column. Approximately 75% of the input biological activity as assayed on Swiss 3T3 cells bound to the column. The column was washed with 0.01 M Tris, pH 7.4 containing 0.15 M NaCl and the proteins eluted with a step gradient of 0.3, 0.6, 1, and 3 M NaCl in the same buffer. The fractions with biological activity eluted after the major protein peak at 0.6 M NaCl (*c*). The fractions after which no optical density 280 (dashed line) was detectable in this step (*c*, fractions 33 through 45) were pooled, acidified, and pumped onto a 250 $\times$ 4.6 mm synchropak RP4 reverse phase column. The proteins were eluted with a linear gradient of 10–80% acetonitrile over 3 h (*d*). Each fraction was assayed for [3 H]-TdR incorporation activity (bar graph) and the optical density was continuously monitored at 212 nm (wavy line). Fractions with the highest biological activity (25–35) were dried under vacuum and run under non-reducing conditions on 15% acrylamide gels containing 0.1% SDS and stained with silver (*e*). There is



no apparent difference between the M_r values of reduced and unreduced protein determined by SDS-PAGE (data not shown). An identical sample in a contiguous lane was cut into 20 equal fractions, the proteins eluted by shaking at 4°C overnight in 0.02% SDS, and the eluant assayed for biological activity on Swiss 3T3 cells. The data are plotted as fold-stimulation of [3 H]-TdR incorporation into Swiss 3T3 cells as a function of fraction (gel slice) number. In this and the other parts of the figure the background (no addition) [3 H]-TdR incorporation was between 500 and 800 c.p.m. for a 5-h pulse of [3 H]-TdR. The remainder (approximately 80%) of the sample shown in *d* was run on an identical SDS-acrylamide gel, transferred electrophoretically to Immobilon and stained with Coomassie blue¹². The indicated areas (*a*–*f*) were cut out for protein microsequencing. The overall yield from 4 litres of medium was about 6 μg or 0.0012% of the starting material. Sequence analysis was carried out on an ABI 470A Protein Sequencer equipped with an on-line ABI 120A PTH Analyzer (Applied Biosystems, Foster City, CA). Yields for PTH amino-acid derivatives were between 33 and 2 picomoles. All fractions which had a readable signal produced the same sequence, and in all cases no signal was detectable in position 7, a probable glycosylation site. X, Unidentified residue. The following amino-acid sequences (single-letter code) were obtained: *a*, VIKPKEXKTEGEKSSE; *b* + *c*, VIKPK-EXKTEGEKSSEK; *d*, VIKPKEXKTEGEK; *e*, VIKP; *f*, XXXPK.

smaller protein containing 78 amino acids, and is 69% homologous to the precursor (Fig. 2c; ref. 4). Both proteins have a motif of six cysteines characteristic of the epidermal growth factor (EGF) family (Fig. 2b). Although there is a strong amino-acid sequence homology between SDGF and amphiregulin, there are a number of differences. The apparent M_r of the most abundant form of SDGF is 10,000 greater than that of amphiregulin. The biological activity of both proteins also seems to be distinct. Amphiregulin inhibits the growth of A431 human carcinoma cells^{5,6} whereas HPLC-purified SDGF has no effect on the growth of A431 cells at concentrations threefold higher than those required for amphiregulin to inhibit growth by 50% (Table 1; ref. 6). Conditioned medium from the amphiregulin-secreting MCF7 cell line is also less potent in stimulating 3T3 cell division than that from JS1 cell cultures (Table 1).

Southern blot analysis of rat genomic DNA was performed to determine whether the SDGF gene cloned is present in the rat genome. Hybridization with SDGF cDNA probe revealed two bands in DNA digested with *Eco*RI and *Pst*I and three in *Hind*III digests (Fig. 3a). Northern blots of SDGF messenger RNA from JS1 cells reveals a principal band of 1.7 kb and minor

bands of 3 and 4.5 kb (arrowheads in Fig. 3c and d). Northern analysis of RNA from several rat tissues shows the strongest hybridization in newborn lung (Fig. 3b, c and d). Weaker reactions requiring longer film exposures were seen with fetal brain, fetal and adult lung, and sciatic nerve (Fig. 3d). No detectable 1.7-kb, 3 or 4.5-kb mRNA bands were seen in adult liver, newborn or adult brain (Fig. 3b, c and d), nor in fetal or adult heart, newborn liver, adult kidney or placenta, or fetal or adult muscle (data not shown). Adult pituitary expresses very low levels of the 1.7-kb mRNA (data not shown).

To determine whether cloned SDGF cDNA exhibits the mitogenic activities found in the JS1 cell culture supernatant, the protein was expressed in human 293 cells. Growth-conditioned medium of transfected 293 cells was mitogenic for Swiss 3T3 cells and primary CNS astrocytes whereas cells transfected with vector alone were inactive (Fig. 1). These observations indicate that the cloned cDNA encodes the heparin-binding mitogen secreted from the JS1 cell line. The purified protein also stimulates DNA synthesis approximately threefold in JS1 cells (Table 1).

The size, biological activities and basic amino-acid

TABLE 1 Stimulation of DNA synthesis and cell division by SDGF

Cell type	Protein	Forskolin	Stimulation	Cell type	Protein	Forskolin	Stimulation
Astrocyte	SDGF (15 pM)		5.2 ± 0.1	Schwann ([³ H]TdR)	TFSup	2 µM	4.6 ± 0.5
	TFSup		4.7 ± 0.3		ContSup		1.1 ± 0.3
	ContSup		1.0 ± 0.1		ContSup	2 µM	6.1 ± 0.8
	GGF (2.5 µg ml ⁻¹)		4.6 ± 0.8		GGF (2.5 µg ml ⁻¹)	—	10 ± 2
	EGF (10 ng ml ⁻¹)		2.6 ± 0.2		GGF (2.5 µg ml ⁻¹)	2 µM	62 ± 5
	TGFα (10 ng ml ⁻¹)		3.2 ± 0.1	A431	SDGF (200 pM)		1.1 ± 0.2
	Serum (2%)		3.1 ± 0.3		JS1 total sup		0.9 ± 0.1
Swiss 3T3	SDGF (15 pM)		23 ± 2		TF Sup		0.9 ± 0.2
	TFSup		35 ± 5		ContSup		0.8 ± 0.3
	ContSup		1.2 ± 0.2		MCF7Sup		0.05 ± 0.01
	GGF (2.5 µg ml ⁻¹)		25 ± 3	JS1	SDGF (15 pM)		3.5 ± 0.3
	EGF (10 ng ml ⁻¹)		30 ± 1		TFSup		2.8 ± 0.3
	TGFα (10 ng ml ⁻¹)		22 ± 3		ContSup		1.2 ± 0.1
	Serum		31 ± 1		Serum		6.1 ± 0.5
Schwann ([³ H]TdR)	MRC7Sup		3.0 ± 0.3	Schwann (cell number)	Nothing		33,900 ± 500
	SDGF (15 pM)		1.2 ± 0.1		ContSup		32,740 ± 1,280
	—	2 µM	5.3 ± 0.7		TFSup		56,700 ± 200
	SDGF (15 pM)	2 µM	5.0 ± 0.5		SDGF (20 pM)		69,800 ± 5,000
	TFSup	—	0.9 ± 0.2		GGF (2.5 µg ml ⁻¹)		138,000 ± 4,000

The culture and measurement of [³H]-TdR incorporation into Swiss 3T3 cells and primary CNS astrocytes was as described in Fig. 1a. Incorporation into JS1 cells was assayed as for 3T3 cells except that they were plated at 1 × 10³ cells per well. Primary Schwann cells were isolated from 2-day neonatal rat sciatic nerve as described by Raff *et al.*⁸ and cultured in DME plus 10% FCS without exposure to partially pure GGF or forskolin. The mitogenic activity of the growth factors was determined by plating the cells on polylysine (5 µg per well) coated 96-well tissue culture microtitre plates in the presence of serum at 1 × 10⁴ cells per well. After 2 days, the culture medium was replaced with N₂ medium¹⁹ and the growth factors added. After 24 h, [³H]-TdR was added and 24 h later, the incorporation of isotope into DNA was determined as with 3T3 cells (Fig. 1). Using a variety of assay conditions (cells on laminin, polylysine and in N₂ or serum-containing medium) no reproducible stimulation of DNA synthesis in primary Schwann cells by SDGF was observed. Partially purified pituitary GGF, a gift from G. Lemke, was always positive. Inhibition of DNA synthesis by amphiregulin in A431 cells was performed as described^{5,6}. Amphiregulin-secreting MCF7 cells were obtained from K. Nikolics (Genentech). Schwann cells were seeded on 35-mm poly-L-lysine coated (50 µg per dish) tissue culture dishes at 2 × 10⁴ cells per dish in N₂ medium plus or minus SDGF. Seven days later the cells were detached into single cells by 10% pancreatin (GIBCO) and counted in a Coulter counter (total cell number per dish). Data are presented as the mean of triplicate determinations ± s.e.m., all done at least three times with the same result. 'Stimulation' is relative to control cultures to which nothing or the appropriate control solvent was added. As amphiregulin inhibited DNA synthesis, a stimulation of 1 indicates no change, 2 is twice the isotope incorporated and 0.05 is 95% inhibition of incorporation. The results are with saturating (maximum stimulation) amounts of material as determined by increasing the growth factor concentration until a saturating (no further effect) response was obtained. Non-responding situations were also tested over a 1,000-fold concentration range. TFSup, Serum-free supernatant of 293 cells transfected with JS1 containing expression plasmid; ContSup, supernatant of cells transfected with plasmid lacking SDGF cDNA; JS1 total sup, JS1 serum-free growth-conditioned medium; MCF7 sup, serum-free growth-conditioned medium from the mammary carcinoma cell line MCF7 stimulated with phorbolsters^{5,6}.

composition of SDGF are also characteristic of a Schwann cell mitogen called glial growth factor (GGF)^{7,8}. In spite of these similarities, SDGF does not stimulate the incorporation of [³H]-TdR into cultured Schwann cells under conditions where a partially purified pituitary fraction containing GGF is active (Table 1). The standard assay used for GGF requires, however, that Schwann cells synthesize DNA between 24 and 48 h after the addition of growth factor. The activity of mitogens that stimulate DNA synthesis later would be missed. To establish whether SDGF is truly mitogenic for cultured Schwann cells, its effect on cell proliferation was determined. Table 1 shows that cells grown in either purified SDGF or in SDGF synthesized

by transfected cells proliferate at a rate about twice that of control cultures over a period of one week. SDGF is therefore a weak mitogen for Schwann cells, an observation compatible with the fact that mRNA for SDGF is present in sciatic nerve (Fig. 3d). As the kinetics of DNA synthesis induced by SDGF and GGF are different, SDGF is not GGF *per se*. SDGF may, however, be a part of the GGF activity. For example, GGF activity from pituitary could contain two or more components of similar *M_r* which copurify. It is therefore possible that SDGF is a component of GGF and that a synergistic interaction with other proteins is required to potentiate the rapid growth of Schwann cells. □

Received 9 July; accepted 3 October 1990.

1. Schubert, D. *et al.* *Nature* **249**, 224–227 (1974).
2. Porter, S., Glaser, L. & Bunge, R. P. *Proc. natn. Acad. Sci. U.S.A.*, **84**, 7768–7772 (1987).
3. Brookes, J. P., Breakefield, X. O. & Martuza, R. L. *Ann. Neurol.* **20**, 317–322 (1986).
4. Plowman, G. D. *et al.* *Molec. cell Biol.* **10**, 1969–1981 (1990).
5. Shoyab, M., Plowman, G. D., McDonald, V. L., Bradley, J. G. & Todaro, G. J. *Science* **243**, 1074–1076 (1989).
6. Shoyab, M., McDonald, V. L., Bradley, J. G. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6528–6532 (1988).
7. Lemke, G. E. & Brookes, J. P. *J. Neurosci.* **4**, 75–83 (1984).
8. Raff, M. C., Abney, E., Brookes, J. P. & Hornby-Smith, A. *Cell* **15**, 813–822 (1978).
9. Baird, A., Schubert, D., Ling, N. & Guillemin, R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2324–2328 (1988).
10. McCarthy, K. D. & deVellis, J. *J. Cell Biol.* **85**, 890–902 (1980).
11. Gorman, C. M. *et al.* *DNA Prot. Engng. Tech.* **2**, 3–10 (1990).
12. Matsudaira, P. *J. Biol. Chem.* **262**, 10035–10042 (1987).
13. Lee, D. C., Rose, T. M., Webb, N. R. & Todaro, G. J. *Nature* **313**, 489–491 (1985).

14. Gray, A., Dull, T. J. & Ullrich, A. *Nature* **303**, 722–725 (1983).
15. Feramisco, J. R., Heffman, D. M., Smart, J. E., Burridge, K. & Thomas, G. P. *J. biol. Chem.* **257**, 11024–11031 (1982).
16. Gubler, U. & Hoffman, B. *J. Gene* **25**, 263–269 (1983).
17. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4767–4771 (1987).
18. Sanger, F., Miklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
19. Bottenstein, J. & Sato, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 514–517 (1979).

ACKNOWLEDGEMENTS. We thank A. Kakizuka, J. Boulter, G. Lemke, G. Weinmaster, K. Umezono and M. LaCorbiere for technical advice and K. C. McFarland and K. Nikolics for preparing some of the serum-free JS1 growth-conditioned medium used in these experiments. We also thank S. Struthers for rat genomic DNA, C. Gorman for the expression vector pCIS2, P. Maher, A. Kakizuka, M. LaCorbiere, and G. Weinmaster for reading the manuscript, and S. Moles for its careful preparation. The JS1 cell line was originally established in culture by J. LeBeau during her thesis work in the laboratory of D.S. H.K. is the recipient of a postdoctoral fellowship from the Human Frontier Science Program, Japan Science Foundation. This work was supported by the National Institutes of Health and the Muscular Dystrophy Association of America.

Mechanosensitive ion channels of *E. coli* activated by amphipaths

Boris Martinac, Julius Adler & Ching Kung

Laboratory of Molecular Biology and the Departments of Genetics and Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, USA

MECHANOSENSITIVE channels have been found in more than 30 cell types, including bacterial^{1,2}, yeast³, plant⁴ and animal cells⁵⁻⁹. Whether tension is transferred to the channel through the lipid bilayer and/or underlying cytoskeleton is not clear. Using the patch-clamp method¹⁰, we found that amphipathic compounds, which are molecules having hydrophobic and hydrophilic character with positive, negative or no net electric charge at pH 7, could slowly activate the mechanosensitive channels of giant *Escherichia coli* spheroplasts, with effectiveness proportional to their lipid solubility. The cationic or anionic amphipaths were able to compensate for each other's effect. After a channel was activated by an amphipath of one charge, if that amphipath was gradually replaced by one with the opposite charge, the channel first inactivated before reactivating. These findings support the view that the mechanical gating force can come from the surrounding lipids.

Cationic amphipaths such as chlorpromazine (CPZ) cause blood cells to form cups, and anionic amphipaths, such as trinitrophenol (TNP), crenate them^{11,12}. According to the bilayer couple hypothesis¹³, these shape changes are the result of preferential insertion of cationic amphipaths into the more negatively charged inner leaflet or anionic species into the less negative

outer leaflet of the membrane lipid bilayer (Fig. 1d)^{14,15}. We used the mechanosensitive channel as a molecular reporter in the *E. coli* outer membrane¹⁶ to detect possible gating forces when amphipaths were applied. The highly asymmetric lipid composition in this membrane^{17,18} leads to charge asymmetry. There are 2×10^6 negative charges in the outer leaflet but more than 3×10^6 negative charges in the inner leaflet of the outer membrane of a single *E. coli*^{18,19}. CPZ (Fig. 1a, c) or TNP (Fig. 1b, c) had no effect on the unit conductance of the mechanosensitive channels in excised inside-out patches, but could reversibly increase their probability of being open to near certainty with half-times of tens of minutes. When amphipaths were removed from the bath (arrows, Fig. 1c), the open probability eventually approached the original low level, which indicates no gross changes in channel or membrane. Channel activation by CPZ or TNP was observed in each of the 13 independent experiments performed. As the time course is slow, several hours of continuous recording are required, and as many patches did not last that long, we present here only the most complete individual experiments.

The open probability of this mechanosensitive channel as a function of suction follows the Boltzmann distribution, in which the mechanical free energy available for gating is linearly dependent on pressure¹. We found that at the drug concentrations applied, the amount of suction necessary for an *e*-fold change in channel activity did not differ significantly in the presence or absence of the drugs, which therefore were having little effect on the mechanosensitivity of the channel itself. Yet either amphipath shifted the Boltzmann curves toward lower suction, which suggests that they lowered the apparent threshold for channel

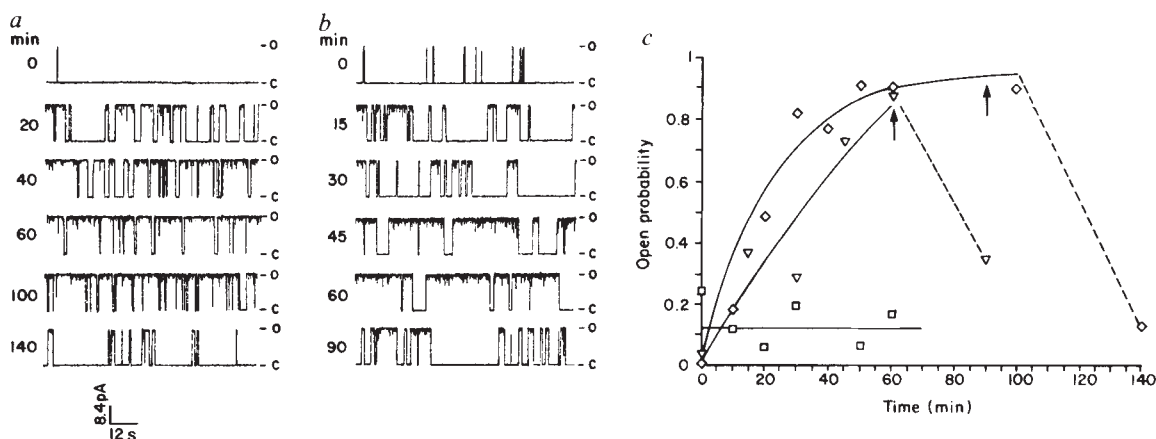
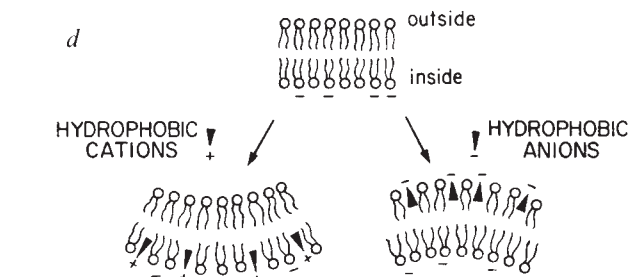


FIG. 1 Experiments showing the effect of CPZ and TNP on the open probability of single mechanosensitive channels over time (a, b, c). Recordings (2 min) of channel activity (a, b) were obtained at marked times during the experiment. Activity was checked every 10 min at the preset suction adjusted to each individual patch. Pipette voltage was +10 mV for all patches. a, With 20 μ M CPZ at 10 mm Hg (1 mm Hg = 133 Pa); b, a different patch in the presence of 500 μ M TNP at 25 mm Hg. Drugs were added at zero time and removed from the bath after 90 (a) and 60 (b) min. 30 min after TNP was washed out, suction was 30 mm Hg. c, Plot of single-channel open probabilities obtained from the current traces of three different patches, like those in a and b, against time after amphipath addition at zero time. Two minutes of recording went into the computation of each datum point. $\square$, No drugs; $\diamond$, 20 μ M CPZ; ∇ , 500 μ M TNP. Arrows indicate the time point when the bath was perfused with the bath solution¹ without drugs. Data points ($\diamond$, ∇) were fitted to single exponential functions given by: $P_o = A(1 - \exp(-t/\tau))$, where P_o is the open probability, t is time, τ is a time constant and A is a fitting parameter. The straight line through the squares should reflect the average open probability of a channel in a control experiment with no drugs present. d, Sketch illustrating the principle of the bilayer couple model after Sheetz and Singer¹³, in which relative monolayer expansions result in contour changes. In reality both leaflets are negatively charged although the inner leaflet is more negatively charged¹⁴.

METHODS. The preparation of giant spheroplasts and solutions as well as



electrical recording were as previously described¹. All recordings were obtained from inside-out excised membrane patches containing at most two active channels. Data were stored on a Gould FM tape recorder (model 6500), digitized off-line at sampling interval of 0.2 ms, and analysed on computer (Indec System, Sunnyvale, CA) with a program developed by Y. Saimi. Open probabilities were calculated by integrating the current through open channels over time of a recording and dividing the integral by the total current through a single channel open over the same period multiplied by the number of active channels. Data points were plotted and fitted to single exponentials by use of the 1-2-3 program (Lotus).

TABLE 1 Amphipathic compounds tested for their effect on the open probability of the mechanosensitive channel

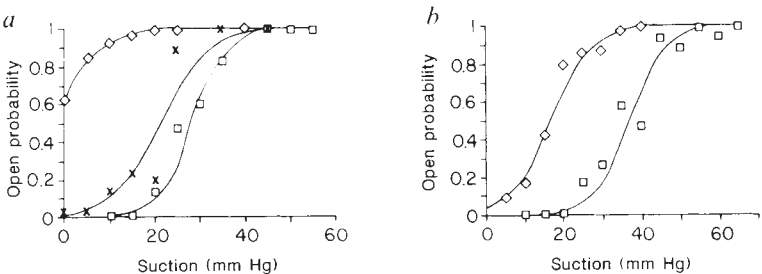
Compound	Ionic form	Concentration tested (M)	$\Delta p_{1/2}$ (mm Hg)	<i>R</i>	Effect on erythrocyte shape ¹²
CPZ	Cationic	2×10^{-5}	42.3	1.0	Cup-former
Procaine	Cationic	5×10^{-4}	17.4	0.41	Cup-former
Tetracaine	Cationic	3×10^{-4}	27.5	0.65	Cup-former
DMAB	Cationic	1×10^{-3}	~0	~0	NT
DEAB	Cationic	1×10^{-3}	11.7	0.28	NT
TNP	Anionic	5×10^{-4}	36.4	0.86	Crenator
Dipyridamole	Anionic	2×10^{-4}	8.1	0.19	Crenator
Lysolecithin	Neutral	2×10^{-4}	21.1	0.50	Crenator

The amount of shift toward lower suction on the pressure axis produced by a drug, $\Delta p_{1/2}$, is used as a quantitative parameter for the drug effect. $\Delta p_{1/2}$ was calculated for each drug as the difference between $p_{1/2}$ values in the absence and the presence of a drug. The $p_{1/2}$ values were obtained from the exponential fits for the effects of drugs on the open probability of the channel over time (see Fig. 1c). Values for open probability were taken at the beginning of an experiment (0 min) and after 60 min in the presence of a drug, and incorporated into the average Boltzmann curve for the mechanosensitive channel (see legend to Fig. 3) with the average value for the *e*-fold change in opening probability of 5.3 ± 1.6 mm Hg (*n* = 13). Thereafter $p_{1/2}$ values correspond to the half-activation points of an average mechanosensitive channel and should reflect the effect of drugs on this average channel. The data are shown from single experiments, in which the largest $\Delta p_{1/2}$ for a particular drug at a given concentration was achieved. The maximal value of $\Delta p_{1/2}$ in this set of experiments was achieved with CPZ. The ratio *R* was calculated by dividing $\Delta p_{1/2}$ for a single drug by the maximal $\Delta p_{1/2}$ to allow a comparison of the drug effects. CPZ, chlorpromazine; DMAB, *N,N*-dimethylaminoethyl benzoate; DEAB, *N,N*-diethylaminoethyl benzoate; TNP, trinitrophenol; lysolecithin, L- α -lysophosphatidylcholine; NT, not tested.

FIG. 2 The effect of 20 μ M CPZ (a) and 500 μ M TNP (b) on the pressure sensitivity of a mechanosensitive channel before perfusion of drug ($\square$), in the presence of indicated drug concentrations for 60 min ($\diamond$), and 40 min after CPZ was washed out of the bath ($\times$) (a); before perfusion of TNP (squares) and in the bath presence of 500 μ M TNP for 60 min (diamonds) (b). Each point represents open probability computed from a 2-min segment of recording as in Fig. 1. The experimental data were fitted to Boltzmann curves given by

$$P_o = \{ \exp [(p - p_{1/2})/s_p] / \{ 1 + \exp [(p - p_{1/2})/s_p] \} ,$$

where P_o is the open probability, p is the applied suction, $p_{1/2}$ is the suction at which channel is open 50% of the time, and s_p is the slope of the plot $\ln [P_o/(1 - P_o)]$ vs suction. The individual slopes are: a, 3.9 ($\square$), 5.4 ($\diamond$) and 5.1 ($\times$) mm Hg, and b, 4.4 ($\square$) and 5.3 ($\diamond$) mm Hg. Linear regression coefficients were in the range between 0.96



and 0.99. Data in a were obtained from the patch in Fig. 1a. Pipette voltage was +10 mV.

activation by exerting an internal gating force on the channel (Fig. 2). Effectiveness in activating the mechanosensitive channel increased with the lipid solubility of the amphipaths (Table 1). We compared the effects of *N,N*-dimethylaminoethyl benzoate (DMAB) with the more hydrophobic *N,N*-diethylaminoethyl benzoate (DEAB). Whereas DEAB could activate the channel within 1 h of recording, DMAB failed to activate in 1 h.

Similarly, the more hydrophobic tetracaine had a lower effective concentration than the less hydrophobic procaine. All drugs, each tested two or more times, took tens of minutes to have effects. This is also consistent with the slow partitioning of the drugs into lipids in contrast to the subsecond direct blockage of Na⁺ channel by amphipathic local anaesthetics²⁰. Cationic amphipaths were more effective than anionic ones, presumably because the latter have to traverse two charged layers of the

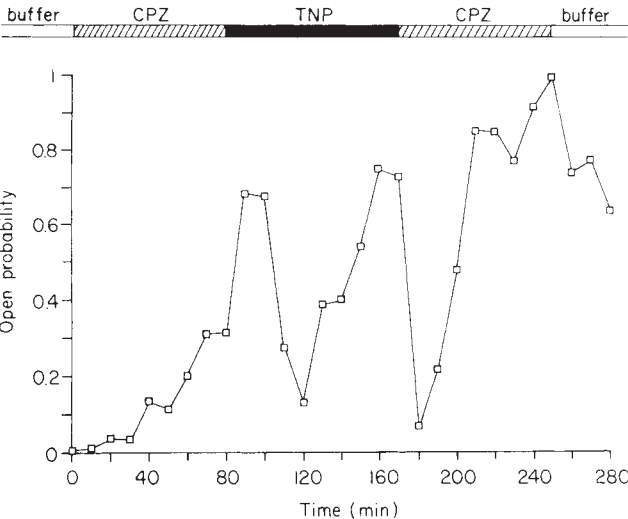


FIG. 3 The compensatory effect of 20 μ M CPZ and 500 μ M TNP (amphipaths with opposite electrical charges) on the open probability of the mechanosensitive channels. Two channels were active in this particular patch. The bath was alternately perfused with buffer¹ containing either CPZ or TNP, or with only buffer at the beginning and end of the experiment. Bars indicate the time point of each solution exchange. Channel activities were recorded every 10 min for 2 min throughout the 280 min of this experiment. The 2-min records all went into the computation of open probability as in Fig. 2. Pipette voltage was +10 mV.

inside-out patches (see Methods in Fig. 1). Lysolecithin, a polar but neutral amphipath, also activated the channel (Table 1).

Amphipaths of opposite charges are known to have compensatory effects on the shape of blood cells^{11,12}. After TNP had been applied to activate the channel, the bathing solution was replaced with CPZ. As replacement gradually occurred in the membrane, the channel was first inactivated and then reactivated (see 170–250 min in Fig. 3). Similarly, after CPZ had been applied to activate the channel, replacing the bathing solution with TNP first deactivated, albeit with a delay, before reactivating (0–160 min in Fig. 3). The activity valleys occurred at times earlier than expected for complete wash-out or wash-in of the two amphipaths. It seems that the balanced presence of both amphipaths in the membrane released the tension, which activates the channel.

A priori, mechanosensitive channels could be gated by forces transferred through the cytoskeleton, whose analogue here is the peptidoglycan to which the outer membrane is attached *in vivo*. Amphipaths would not be expected to associate strongly with peptidoglycan, especially not with strengths correlated with their lipid solubility. *E. coli* mechanosensitive channels have been successfully reassembled from purified membranes onto azolectin liposomes^{21,22}. Unless the peptidoglycan copurifies with the channel and reattaches properly upon reassembly, it cannot be the element that transfers the gating force. It seems to us that the force gating the mechanosensitive channels is from the lipid bilayer and the amphipaths probably generate this force by differential insertion into the two leaflets. Alamethicin, which in pure lipid bilayers assembles into multisubunit ion channels, was also shown to express mechanosensitivity when pressure was applied to bilayer patches²³.

We can envisage the amphipath's increasing the membrane tension, T , which gates the mechanosensitive-channel, by increasing the radius, r , of the patch calotte under a fixed applied pressure, p , as governed by Laplace's law, $T = rp/2$. The fact that a channel can be activated by amphipaths of either charge echoes the following fact: either suction or pressure activates the mechanosensitive channels of *E. coli*, yeast and frog lens epithelium^{1,3,24}. This is analogous to a grommet on a sail, where the grommet is stretched according to wind strength (force) but not according to which way the sail bows (direction). Amphipaths could be used to understand further the physics of mechanosensitive channels and the biology of mechanosensation. □

A mutant T4 lysozyme displays five different crystal conformations

H. R. Faber & B. W. Matthews

Institute of Molecular Biology, Department of Physics and Howard Hughes Medical Institute, University of Oregon, Eugene, Oregon 97403, USA

PHAGE T4 lysozyme consists of two domains between which is formed the active-site cleft of the enzyme^{1,2}. The crystallographically determined thermal displacement parameters for the protein² suggested that the amino terminal of the two domains undergoes 'hinge-bending' motion about an axis passing through the waist of the molecule. Such conformational mobility may be important in allowing access of substrates to the active site of the enzyme¹. We report here a crystallographic study of a mutant T4 lysozyme which demonstrates further the conformational flexibility of the protein. A mutant form of the enzyme with a methionine residue (Met 6) replaced by isoleucine crystallizes with four independent molecules in the crystal lattice. These four molecules have distinctly different conformations. The mutant protein can also crystallize in standard form with a structure very similar to the wild-type protein. Thus the mutant protein can adopt five different crystal conformations. The isoleucine for methionine substitution at the intersection of the two domains of T4 lysozyme apparently enhances the hinge-bending motion presumed to occur in the wild-type protein, without significantly affecting the catalytic activity or thermal stability of the protein.

Figure 1 shows the structure of T4 lysozyme^{1,2}, colour-coded to show the 'thermal motion' displayed in crystals of the wild-type protein². Analysis of the crystallographically determined thermal displacement parameters suggested that hinge-bending motion of the lower, N-terminal domain probably occurs in solution.

M6I is one of a number of randomly generated temperature-sensitive mutants of T4 lysozyme³. The activity of the purified^{4,5} protein, as measured by the rate of hydrolysis of cell walls⁶, seems identical with that of wild-type lysozyme. The melting temperature of the protein⁷ at pH 6.0 is 3.4 °C lower than that of wild type, corresponding to a reduction of 1.4 kcal mol⁻¹ in the free energy of stabilization⁷.

M6I, set up to crystallize under the standard conditions², gives two types of crystals. The first (form I) is isomorphous with the trigonal crystals of wild-type lysozyme. The structure of M6I in this form was determined and refined⁸ using oscillation diffraction data⁹ to 1.8 Å resolution (Table 1). Its structure is very similar to wild type (additional details will be given elsewhere).

The second crystal modification¹⁰ (form II) is orthorhombic, with four molecules in the asymmetric unit (Table 1). To solve the orthorhombic crystal structure it was necessary not only to use isomorphous replacement to determine the locations of the four independent molecules in the crystal lattice, but also to separately place the respective N-terminal and C-terminal domains of each of these four lysozyme molecules. A heavy atom derivative was first obtained by soaking the crystals in 2.5 × 10⁻⁴ M HgCl₂ for two days (compare ref. 1), and a difference Patterson map calculated using diffractometer data to 5 Å resolution (Table 1) clearly indicated that there were four sites of mercury substitution. Using these sites, a single isomorphous replacement (SIR) electron-density map at 5 Å resolution was calculated which had a number of features that appeared to correspond to α helices. Inspection¹¹ of this low-resolution map, in the light of the known structure of wild-type T4 lysozyme, suggested the locations of two lysozyme molecules. In each of these molecules, Cys 54 was close to a mercury-binding site, adding credence to the interpretation. It was not

Received 26 April; accepted 14 September 1990.

- Martinac, B., Buechner, M., Delcour, A. H., Adler, J. & Kung, C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2297–2301 (1987).
- Zoratti, M. & Petronilli, V. *FEBS Lett.* **240**, 105–109 (1988).
- Gustin, M. C., Zhou, X.-L., Martinac, B. & Kung, C. *Science* **242**, 762–765 (1988).
- Falke, L. C., Edwards, K. L., Pickard, B. G. & Misler, S. *FEBS Lett.* **237**, 141–144 (1988).
- Guharay, F. & Sachs, F. *J. Physiol., Lond.* **352**, 685–701 (1984).
- Sigurdson, W. J., Morris, C. E., Brezden, B. L. & Gardner, D. R. *J. exp. Biol.* **127**, 191–209 (1987).
- Hudspeth, A. J. *Hearing Res.* **22**, 21–27 (1986).
- Ohmori, H. *J. Physiol., Lond.* **350**, 561–581 (1984).
- Kirber, M. T., Singer, J. J. & Walsh, J. V., Jr. *Biophys. J.* **51**, 252a (1987).
- Hamill, O. P., Marty, Y., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Sheetz, M. P. & Singer, S. J. *J. Cell Biol.* **70**, 193–203 (1976).
- Deuticke, B. *Biochim. biophys. Acta* **163**, 494–500 (1968).
- Sheetz, M. P. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4457–4461 (1974).
- Rothman, J. E. & Lenard, J. *Science* **195**, 743–753 (1977).
- Verkley, A. J. et al. *Biochim. biophys. Acta* **323**, 178–193 (1973).
- Buechner, M., Delcour, A. H., Martinac, B., Adler, J. & Kung, C. *Biochim. biophys. Acta* **1024**, 111–121 (1990).
- Raetz, C. R. A. *Rev. Genet.* **20**, 253–295 (1986).
- Nikaido, H. & Vaara, M. in *Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology* Vol. 1 (eds Neidhardt, F. C. et al.) 7–15 (Am. Soc. Microbiol., Washington DC, 1987).
- Lugtenberg, B. & van Alphen L. *Biochim. biophys. Acta* **737**, 51–115 (1983).
- Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA, 1984).
- Delcour, A. H., Martinac, B., Adler, J. & Kung, C. *Biophys. J.* **56**, 631–636 (1989).
- Berrier, C., Coulombe, A., Houssin, C. & Ghazi, A. *FEBS Lett.* **259**, 27–32 (1989).
- Opsahl, L. R., Mak, D. D. & Webb, W. W. *Biophys. J.* **57**, 321a (1990).
- Cooper, K. E., Tang, J. M., Rae, J. L. & Eisenberg, R. S. *J. memb. Biol.* **93**, 259–269 (1986).

ACKNOWLEDGEMENTS. We thank C. Hirschler, L. Olds and P. Hanson for technical assistance; members of our laboratories for comments on the manuscript; and D. L. Nelson for discussion and for DMAB and DEAB, which were synthesized by J. L. Browning. Supported by the NIH and the Markey Trust.

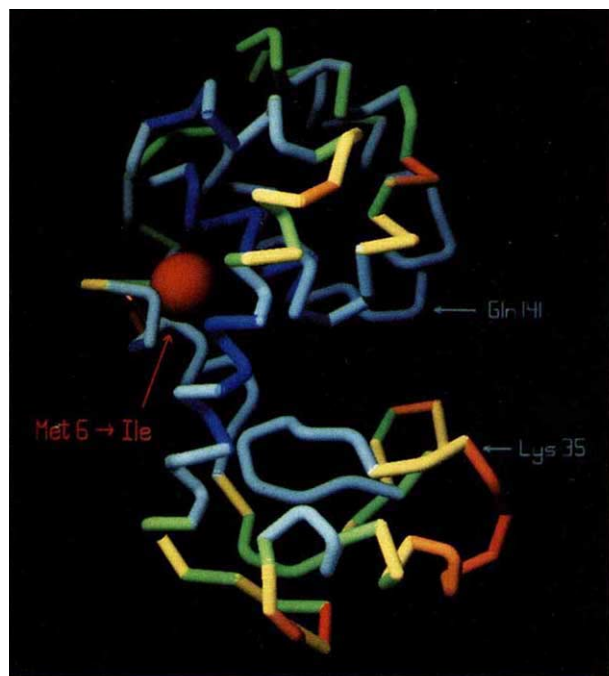


FIG. 1 Backbone of wild-type T4 lysozyme colour-coded according to the thermal parameters of the backbone atoms in the refined crystal structure². Dark blue indicates regions of least thermal motion; red indicates regions of largest thermal parameters. The lower (N-terminal) domain has few contacts with neighbouring molecules in the crystal and, as can be seen from the gradation of colour (see also ref. 2), displays pronounced hinge-bending displacement. The orange sphere shows the location of the side chain of Met 6. In the different crystal structures of the Met 6 $\rightarrow$ Ile variant the α carbons of Lys 35 and Gln 141 move apart by up to 8.6 Å.

possible, however, to locate the remaining two molecules by visual inspection. Therefore a series of six-dimensional searches was performed in which the backbone atoms of wild-type lysozyme were rotated and translated through the SIR map. Agreement with the map was measured by summing the electron density at the backbone ($C\alpha$, N, C, O) atoms. To reduce the search time it was assumed that each of the four lysozyme molecules in the crystal would have Cys 54 or Cys 97 in the vicinity of one of the four known mercury-binding sites. This was consistent with the original determination of the structure of wild-type lysozyme¹. The search function had four very strong peaks, 8.5–12.0 s.d. above background, clearly indicating the positions of four lysozyme molecules. Two of these locations corresponded to those that had been determined by visual inspection. Notwithstanding this clear evidence for the molecular locations, a model with four wild-type lysozyme

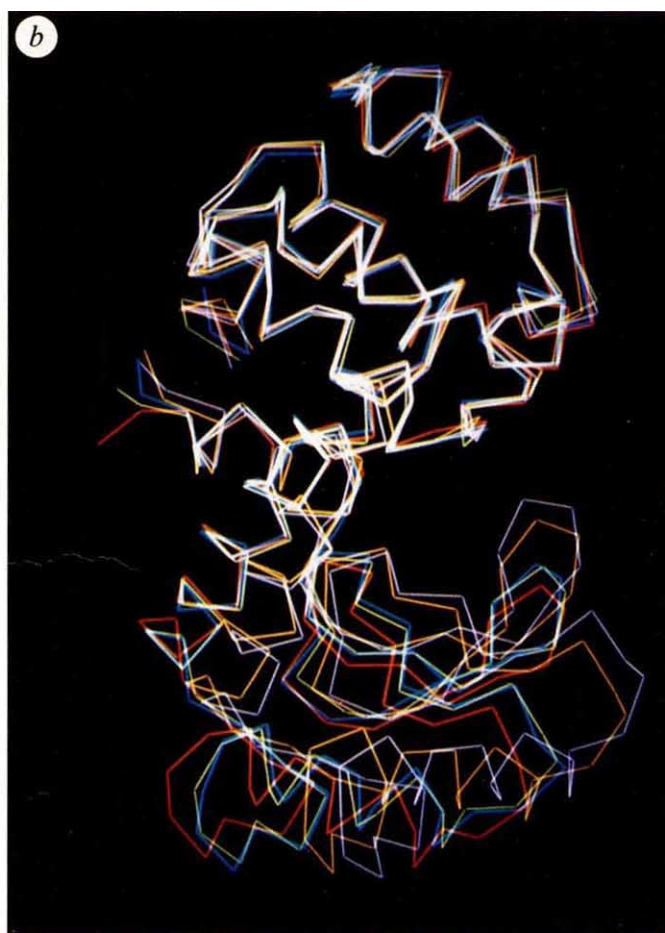
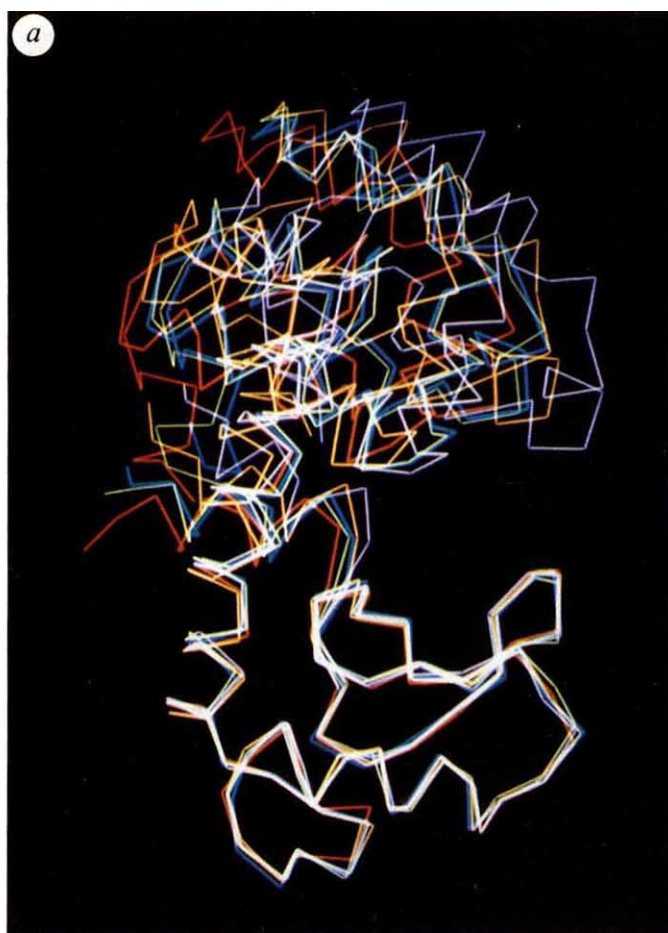


FIG. 2 Superposition of the backbone of five different crystal structures of the mutant lysozyme M6I. The purple molecule is from the wild-type crystal form and has the smallest hinge-bending angle. *a*, Molecules superimposed

according to their N-terminal domains (residues 15–60); *b*, molecules superimposed according to their C-terminal domains (residues 80–160).

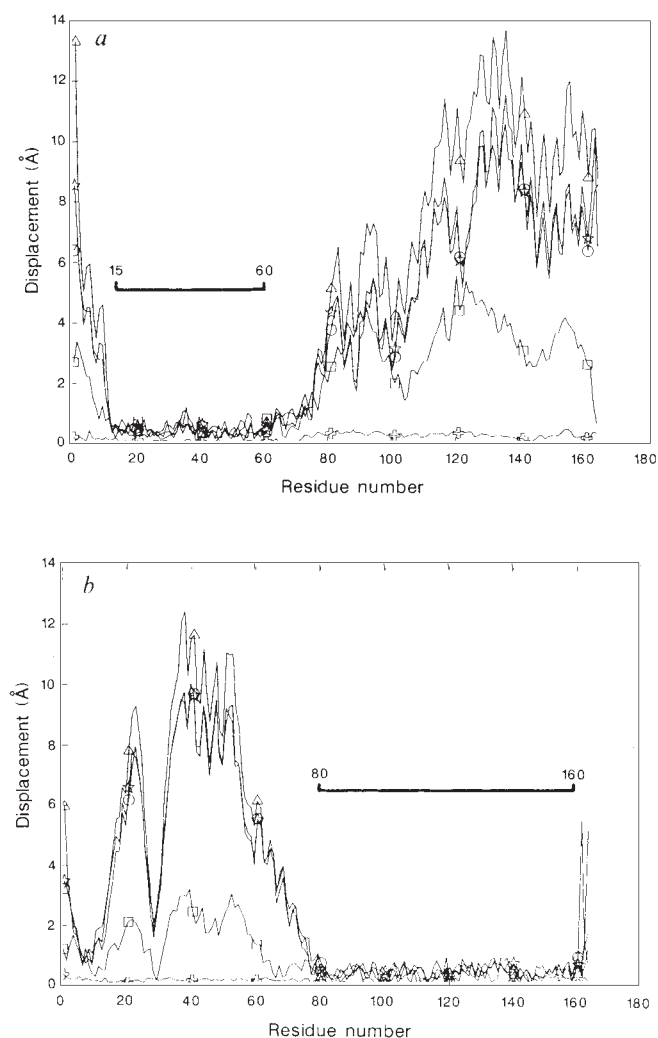


FIG. 3 Differences in the backbone conformation of M6I lysozymes. The four M6I lysozymes from the form II crystal form (square, circle, star, triangle), and the M6I lysozyme from the form I crystals (cross) superimposed on the backbone of wild-type lysozyme. The ordinate shows the distance between corresponding α carbon atoms. In *a*, the superposition is based on the α carbon atoms in the N-terminal domain (residues 15–60; solid line). In *b*, the superposition is based on α carbons 80–160 from the C-terminal domain, also indicated by a solid line.

FIG. 4 Difference-distance matrix showing the difference in C α –C α distances for two of the M6I molecules in the form II crystal form relative to the M6I molecule in the form I crystals (coloured purple in Fig. 2). The top right half of the figure compares the red and the purple molecules (that is, the most open and most closed molecules); the bottom left half of the figure compares the yellow and purple molecules (the two most closed). The function plotted (for the top right half of the figure, for example) is

$$\Delta r_{ij} = (r_{ij})_{\text{red}} - (r_{ij})_{\text{purple}}$$

where r_{ij} is the distance between the i th and j th α -carbon atoms in the specified M6I lysozyme molecule. Contours are drawn at 1, 2, 3 Å... (solid) and -1, -2, -3 Å... (dashed). The locations of the α helices and β sheets are indicated by the letters α and β . Because the function Δr_{ij} has a similar overall appearance in the bottom left and top right halves of the figure, but differs in scale, it shows that the hinge-bending displacement is similar (although not identical) in each case, but of different magnitude.

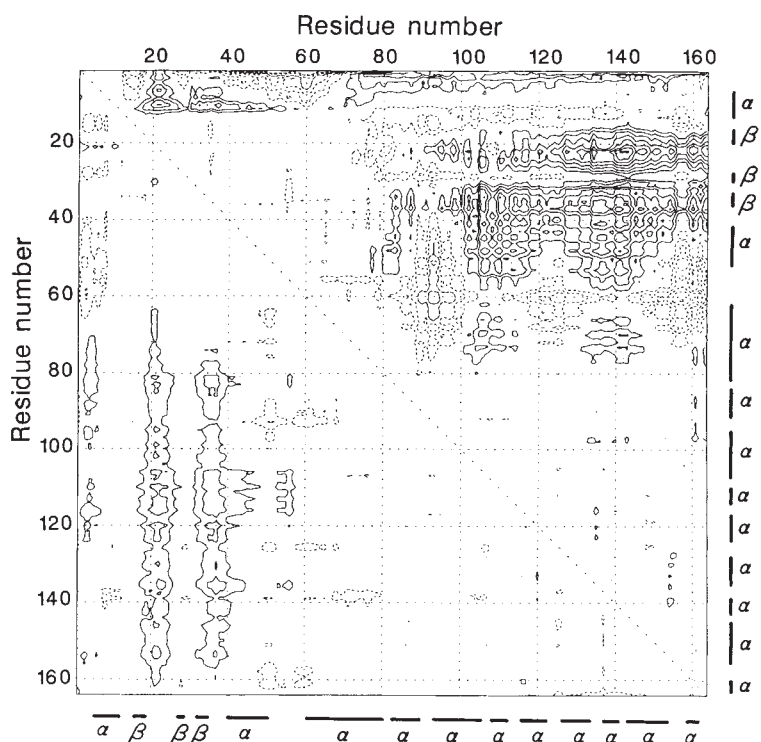


TABLE 1 Crystallographic data

Modification	Form I	Form II
Crystals		
Space group	$P3_121$	$P2_12_12_1$
Cell dimensions (a, b, c) (Å)	61.1, 61.1, 97.6	72.2, 73.8, 150.5
Molecules per asymmetric unit	1	4
Volume of crystals occupied by solvent (%)	56	54
Data collection		
Diffractometer data	—	M6I HgCl ₂
Resolution (Å)	—	5.0
Reflections	—	3,526
R_{sym} (%)	—	6.6
Film data		M6I
Resolution (Å)	1.8	2.1
Reflections	13,645	28,684
R_{merge} (%)	5.2	12.2
Refinement		
Residual (%)	15.5	23.6
No. solvent molecules	153	96
r.m.s. deviations from ideal values		
Bonds (Å)	0.013	0.013
Bond angles (degrees)	2.1	2.7
Planarity (Å)	0.014	0.014

molecules placed at the indicated sites did not refine⁸ satisfactorily. Refinement in which each domain was treated as a separate unit also did not converge satisfactorily. The answer to this dilemma was found by repeating the six-dimensional rotation/translation searches in the original SIR electron density map, not with the whole molecule but with the N- and C-terminal domains separately. This required a total of eight searches, each of which gave a very clear peak 10.1–15.8 s.d. above the average value. This model, in which the N- and C-terminal domain of each lysozyme molecule has been located independently, had an initial R -value of 44.4% at 5 Å resolution and refined⁸ smoothly to the current residual of 23.6% at 2.1 Å resolution (Table 1).

The striking result is that each of the four independent M6I lysozyme molecules in the form II crystals has a different hinge-bending angle between the upper and lower domains. Furthermore, each of these conformations is different from the M6I structure as seen in the form I (wild-type) crystal modification. The relationship between the five different M6I structures is illustrated in Fig. 2. Considered alone, the structures of the C-terminal domains of the different allomorphs are virtually superimposable (Figs 2a and 3a). Similarly, the N-terminal domains are also extremely similar to each other (Figs 2b and 3b). The difference between the five different M6I structures is in the hinge-bending angle between the domains. Relative to M6I in the wild-type crystal form, the hinge-bending angle increases by 9.1°, 25.6°, 27.4° and 32.3° in the four allomorphs in the form II crystal.

This behaviour can be shown quantitatively by using a 'different-distance' matrix^{12,13} (Fig. 4). Such a plot shows the change in the distance between corresponding pairs of α -carbon atoms in two different M6I molecules. Areas of Fig. 4 that are featureless indicate regions of three-dimensional structure that are essentially identical in the two structures being compared. Residues 12–76 constitute one such domain. This corresponds to the lower, or N-terminal, domain (Figs 1 and 2) as well as most of the long helix that connects the lower and upper domain. Residues 77–164, corresponding to the upper domain, also clearly form an essentially rigid unit. The α helix of residues 3–11, which is located in the 'waist' of the molecule (Fig. 1) and includes the site of the Met 6→Ile replacement, seems to be linked more closely to the upper rather than the lower domain. Figure 4 also shows those regions of the two structures that have the greatest relative displacement. Lys 35 and Gln 141 for example, which are on opposite sides of the active site cleft (Fig. 1), increase their separation by 8.6 Å in the most open M6I structure relative to the most closed.

Inspection of the difference shift plots for the other pairs of M6I structures shows that they are similar (but not identical) in appearance, except that the absolute magnitude of the shifts varies from case to case (Fig. 4). Thus the five different M6I allomorphs seem to sample a continuous range of conformations that could occur in solution. The refined structures of the M6I allomorphs in the immediate vicinity of residue 6 seem to be very similar (atoms within a 5 Å radius are superimposable within ~0.25 Å). Therefore the locus of the hinge-bending is not at residue 6 itself. Rather, there are three places in the molecule at which much of the hinge-bending seems to be localized. The most obvious is at or near Ile 13, which is within an extended segment connecting the upper and lower domains (Fig. 1). The other two places are, respectively, in the vicinity of residues 58–60 and 79–80, that is, at the ends of the long α helix connecting the upper and lower lobe of the lysozyme molecule (Fig. 1). At each of these three sites the hinge-bending motion does not seem to be generated by rotation about a single bond but rather by a combination of backbone conformational changes extending over several residues. Thus, although the global changes observed in the different M6I molecules are large, they are effected by quite small changes at the sites of hinge-bending. This suggests that hinge-bending motion in M6I lysozyme is a low-energy deformation. For each of the five M6I structures, the thermal factors of the backbone atoms within the hinge-bending regions are close to the average value for whole structure, indicating that these atoms are not especially mobile or disordered.

Large conformational changes in proteins are known to occur in response to the binding of ligands (see for example refs 14–17, 21). Also, examples are known of conformational variability in proteins such as immunoglobulins that consist of essentially independent domains (refs 17–19). The present example seems to be the first case in which a single macromolecular species, which folds as a cooperative unit^{7,20}, displays large-scale conformational variability. It provides a snapshot of the conformational flexibility accessible to macromolecules. □

Received 11 May; accepted 3 September 1990.

- Matthews, B. W. & Remington, S. J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4178–4182 (1974).
- Weaver, L. H. & Matthews, B. W. *J. molec. Biol.* **193**, 189–199 (1987).
- Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C. & Matthews, B. W. *Biochemistry* **26**, 3754–3758 (1987).
- Alber, T. & Matthews, B. W. *Meth. Enzym.* **154**, 511–533 (1987).
- Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E. & Dahlquist, F. W. *Meth. Enzym.* **177**, 44–73 (1989).
- Tsugita, A., Inouye, M., Terzaghi, E. & Streisinger, G. *J. biol. Chem.* **243**, 391–397 (1968).
- Becktel, W. J. & Schellman, J. A. *Biopolymers* **26**, 1859–1877 (1987).
- Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489–501 (1987).
- Schmid, M. F. et al. *Acta crystallogr.* **A37**, 701–710 (1981).
- Brennan, R. G., Wozniak, J., Faber, R. & Matthews, B. W. *J. Crystal Growth* **90**, 160–167 (1988).
- Jones, T. A. in *Crystallographic Computing* (ed. Sayre, D.) 303–317 (Oxford University Press, 1982).

- Phillips, D. C. *Biochem. Soc. Symp.* **31**, 11 (1970).
- Nishikawa, K., Ooi, T., Isogai, Y. & Saito, N. *J. Phys. Soc. Jpn.* **32**, 1331–1337 (1972).
- Baldwin, J. & Chothia, C. *J. molec. Biol.* **129**, 175–220 (1979).
- Steitz, T. A., Shoham, M. & Bennett, W. S., Jr. *Phil. Trans. R. Soc.* **B293**, 43–52 (1981).
- Remington, S. J., Wiegand, G. & Huber, R. *J. molec. Biol.* **158**, 111–152 (1982).
- Huber, R. & Bennett, W. S. Jr. *Biopolymers* **22**, 261–279 (1983).
- Colman, P. M., Deisenhofer, J., Huber, R. & Palm, W. *J. molec. Biol.* **100**, 257–282 (1976).
- Huber, R., Deisenhofer, J., Colman, P. M., Matsushima, M. & Palm, W. *Nature* **264**, 415–420 (1976).
- Hawkes, R., Grütter, M. G. & Schellman, J. *J. molec. Biol.* **175**, 195–212 (1984).
- Anderson, B. F., Baker, H. M., Norris, G. E., Rumball, S. V. & Baker, E. N. *Nature* **344**, 784–787 (1990).

ACKNOWLEDGEMENTS. We thank J. Wozniak for help in purifying and crystallizing the mutant lysozyme and S. Roderick, S. J. Remington, L. Weaver and D. Tronrud for help with data collection and computation. This work was supported in part by the National Institutes of Health and the Markey Charitable Trust.